

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 1999, 13:01:00 ; Search time 40.75 Seconds
(without alignments)
46,500 Million cell updates/sec

Title: US-09-142-524-1

Perfect score: 406
Sequence: 1 MKYTVAFNQGPRKRVFIKR.....IASRRVDGIIAAYQNPASWK 80

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	406	100.0	80	1	W27369	Multi-epitope pept
2	327.5	80.7	105	1	W27370	Multi-epitope pept
3	320	78.8	134	1	W27371	Multi-epitope pept
4	144	35.5	47	1	W80353	Sugi allergen prot
5	135	33.3	81	1	W80358	Sugi allergen prot
6	125.5	30.9	47	1	W80352	Sugi allergen prot
7	121.5	29.9	47	1	W80351	Sugi allergen prot
8	120	29.6	47	1	W80357	Sugi allergen prot
9	114	28.1	47	1	W80356	Sugi allergen prot
10	99	24.4	514	1	R53690	Japanese cedar pol
11	99	24.4	514	1	R69792	Japanese cedar pol
12	99	24.4	460	1	R69791	Japanese cedar pol
13	99	24.4	514	1	R74333	Japanese cedar pol
14	99	24.4	514	1	R81586	Japanese cedar pol
15	99	24.4	514	1	R93599	Japanese cedar pol
16	98	24.4	35	1	W80342	Sugi allergen prot
17	94.5	23.3	33	1	W80340	Sugi allergen prot
18	92.5	22.8	33	1	W80339	Sugi allergen prot
19	91	22.4	33	1	W80338	Sugi allergen prot
20	89	21.9	17	1	R81582	Sugi allergen prot
21	88	21.7	17	1	W80347	Sugi allergen prot
22	88	21.7	47	1	W80350	Sugi allergen prot
23	83	20.4	17	1	R81580	Sugi allergen prot
24	83	20.4	17	1	W80345	Sugi allergen prot
25	81	20.0	15	1	R97884	Japanese cedar pol
26	81	20.0	15	1	W57760	Japanese cedar pol
27	79	19.5	15	1	R97908	Japanese cedar pol
28	79	19.5	15	1	W57764	Japanese cedar pol
29	74	18.2	514	1	W04346	Chamaecyparis obtu
30	74	18.2	354	1	W04344	Chamaecyparis obtu
31	74	18.2	375	1	W04345	Chamaecyparis obtu
32	74	18.2	354	1	W42121	Japanese cypress p
33	74	18.2	514	1	W42122	Japanese cypress p
34	73	18.0	30	1	R45588	Japanese cypress p
35	73	18.0	367	1	R45577	Japanese cypress p
36	73	18.0	15	1	R97907	Japanese cypress p
37	73	18.0	15	1	W57763	Japanese cypress p
38	71	17.5	374	1	R45586	Japanese cypress p
39	71	17.5	36	1	R45587	Japanese cypress p
40	71	17.5	30	1	R45587	Japanese cypress p
41	71	17.5	374	1	R45541	Japanese cypress p
42	71	17.5	374	1	R60166	Japanese cypress p
43	71	17.5	353	1	R75388	Japanese cypress p

44 71 17.5 374 1 R82490
45 71 17.5 353 1 R81587

Cry j I Japanese C
Cedar pollen aller

ALIGNMENTS

RESULT 1
W27369
W27369: standard; peptide; 80 AA.
AC W27369:
DE 24-MAR-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #1.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
PN W09732600-AA.
PD 12-SEP-1997.
PE 10-MAR-1997: J00740.
PR 10-MAR-1996: JP-080702.
PA (MEIP) MEIJI MILK PROD CO LTD.
PI Dairiki K, Iwama A, Kuno K, Kume A, Some T;
DR WPI: 97-470495/43.
PT Peptide immuno:therapeutic agent to treat allergic diseases -
PT contains multi-epitope peptide containing T cell epitope regions
PT from different allergens
PS Claim 6; Page 31; 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions: have a positively index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 80 AA:

Query Match 100.0%; Score 406; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 4.2e-48;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKYTVAFNQGPRKRVFIKRVSVIIHGRRIDIFASKNFHLQNTIGTGRISIKITSGK 60
DB 1 MKYTVAFNQGPRKRVFIKRVSVIIHGRRIDIFASKNFHLQNTIGTGRISIKITSGK 60

OY 61 IASRRVDGIIAAYQNPASWK 80
DB 61 IASRRVDGIIAAYQNPASWK 80

RESULT 2
W27370
W27370: standard; peptide; 105 AA.
AC W27370:
DE 24-MAR-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #2.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
PN W09732600-AA.
PD 12-SEP-1997.
PE 10-MAR-1997: J00740.
PR 10-MAR-1996: JP-080702.
PA (MEIP) MEIJI MILK PROD CO LTD.
PI Dairiki K, Iwama A, Kuno K, Kume A, Some T;
DR WPI: 97-470495/43.
PT Peptide immuno:therapeutic agent to treat allergic diseases -
PT contains multi-epitope peptide containing T cell epitope regions
PT from different allergens
PS Claim 6; Page 31; 58pp; Japanese.

CC The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 SQ Sequence 105 AA;

Query Match 80.7%; Score 327.5; DB 1; Length 105;
 Best Local Similarity 67.6%; Pred. No. 2.8e-37;
 Matches 71; Conservative 1; Mismatches 8; Indels 25; Gaps 2;

QY 1 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGTGRIS----- 54
 DB 1 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGTGRKKNRIMLQ 60
 QY 54 -LKLTSGKIA-----SRVDGIIAAYONPASM 80
 61 FAKLTGFTIMGRRLKMPYIAGYKTFDGRVDTGIIAAYONPASM 105

RESULT 3
 W27371
 ID W27371 standard; peptide; 134 AA.
 AC W27371;
 DT 24-MAR-1998 (first entry)
 DE Multi-epitope peptide used as immunotherapeutic agent #3.
 KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 OS Synthetic.
 PN W09732600-AA.
 PF 12-SEP-1997;
 PR 10-MAR-1997; J00740.
 PR 10-MAR-1996; JP-080702.
 PA (MEIP) MEIJI MILK PROD CO LTD.
 PI Dairiki K, Iwama A, Kume A, Sone T;
 DR WPI; 97-470495/43.
 PT Peptide immuno: therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PS from different allergens
 PS Claim 6; Page 32; 58pp; Japanese.

CC The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 SQ Sequence 134 AA;

Query Match 78.8%; Score 320; DB 1; Length 134;
 Best Local Similarity 53.7%; Pred. No. 4e-36;
 Matches 72; Conservative 2; Mismatches 6; Indels 54; Gaps 2;

QY 1 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGTGR----- 52
 DB 1 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGTGRKKNRIMLQ 60
 QY 52 -----ISLKT-----TSKIASRV 66
 DB 61 FAKLTGFTIMGRRLKMPYIAGYKTFDGRVDTGIIAAYONPASM 120
 QY 67 DGIITAYONPASM 80
 |||||||||||||

DB 121 DGIITAYONPASM 134

RESULT 4

ID W80353 standard; peptide; 47 AA.
 AC W80353;
 DT 11-JAN-1999 (first entry)
 DE Sugl allergen protein Cryj1 derived epitope for T cells.
 KW T cell epitope; sugl allergen proteins Cryj1; Cryj2; treatment;
 KW sugl-pollinosis; allergic reaction; pollen.
 OS Synthetic.
 PN J10259198-A.
 PF 29-SEP-1998;
 PR 22-DEC-1997; 353448.
 PR 24-DEC-1996; JP-343441.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 DR WPI; 98-577037/49.
 PT A linked T cell epitope peptide - used for the treatment of
 PT sugl-pollinosis
 PS Claim 10; Page 5; 21pp; Japanese.
 CC W80353-58 represent epitopes for T cells, derived from the sugl allergen
 CC proteins Cryj1 (W80353-44, W80350-53 and W80356-58) and Cryj2 (W80345-49
 CC and W80354-55). The peptides are useful for the treatment of
 CC sugl-pollinosis, an allergic reaction of the body to pollen.
 SQ Sequence 47 AA;

Query Match 35.5%; Score 144; DB 1; Length 47;
 Best Local Similarity 44.3%; Pred. No. 8.9e-13;
 Matches 35; Conservative 0; Mismatches 0; Indels 44; Gaps 2;

QY 1 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGTGRISLKTSGK 60
 DB 13 MKVTAFNFG-----FASKNFHLQKNT----- 36
 QY 61 IASRRVDGIIAAYONPASM 79
 DB 36 -----GIIAAYONPASM 47

RESULT 5
 W80358
 ID W80358 standard; peptide; 81 AA.
 AC W80358;
 DT 11-JAN-1999 (first entry)
 DE Sugl allergen protein Cryj1 derived epitope for T cells.
 KW T cell epitope; sugl allergen proteins Cryj1; Cryj2; treatment;
 KW sugl-pollinosis; allergic reaction; pollen.
 OS Synthetic.
 PN J10259198-A.
 PF 29-SEP-1998;
 PR 22-DEC-1997; 353448.
 PR 24-DEC-1996; JP-343441.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 DR WPI; 98-577037/49.
 PT A linked T cell epitope peptide - used for the treatment of
 PT sugl-pollinosis
 PS Claim 10; Page 5; 21pp; Japanese.
 CC W80353-58 represent epitopes for T cells, derived from the sugl allergen
 CC proteins Cryj1 (W80353-44, W80350-53 and W80356-58) and Cryj2 (W80345-49
 CC and W80354-55). The peptides are useful for the treatment of
 CC sugl-pollinosis, an allergic reaction of the body to pollen.
 SQ Sequence 81 AA;

Query Match 33.3%; Score 135; DB 1; Length 81;
 Best Local Similarity 46.7%; Pred. No. 3e-11;
 Matches 35; Conservative 2; Mismatches 8; Indels 30; Gaps 2;

QY 1 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGTGRISLKTSGK 60

DB 14 MKVTAFNFGP-----FASKNFHLOKNT-----KLTSKG 43
 OY 61 IASRRVDGIIAYON 75
 DB 44 IASCLNGLVHANN 58

RESULT 6

W80352 standard; peptide; 47 AA.
 ID W80352
 AC W80352:
 DT 11-JAN-1999 (first entry)
 DE Sugi allergen protein Cryj1 derived epitope for T cells.
 KW T cell epitope; sugi allergen proteins Cryj1; Cryj2; treatment;
 KM sugi-pollinosis; allergic reaction; pollen.
 OS Synthetic.
 PN J10259198-A.
 PD 29-SEP-1998.
 PF 22-DEC-1997: JP-343441.
 PR 24-DEC-1996: JP-343441.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 DR WPI: 98-577037/49.
 PT A linked T cell epitope peptide - used for the treatment of
 PT sugi-pollinosis
 PS Claim 10; Page 5; 21pp; Japanese.
 CC W80339-58 represent epitopes for T cells, derived from the sugi allergen
 CC proteins Cryj1 (W80339-44, W80350-53 and W80356-58) and Cryj2 (W80345-49
 CC and W80354-55). The peptides are useful for the treatment of
 CC sugi-pollinosis, an allergic reaction of the body to pollen.
 SQ Sequence 47 AA;

Query Match 30.9%; Score 125.5; DB 1; Length 47;
 Best Local Similarity 50.8%; Pred. No. 2.9e-10; Indels 31; Gaps 2;
 Matches 32; Conservative 0; Mismatches 0; Indels 31; Gaps 2;
 OY 1 MKVTAFNFGPNNRVRIFIKRVSNVLIHGRRIDFASKNFHLOKNTIGTGRIISLKLTSKG 60
 DB 13 MKVTAFNFGP-----FASKNFHLOKNT-----KLTSKG 41
 OY 61 IAS 63
 DB 42 IAS 44

RESULT 7

W80351 standard; peptide; 47 AA.
 ID W80351
 AC W80351:
 DT 11-JAN-1999 (first entry)
 DE Sugi allergen protein Cryj1 derived epitope for T cells.
 KW T cell epitope; sugi allergen proteins Cryj1; Cryj2; treatment;
 KM sugi-pollinosis; allergic reaction; pollen.
 OS Synthetic.
 PN J10259198-A.
 PD 29-SEP-1998.
 PF 22-DEC-1997: JP-343441.
 PR 24-DEC-1996: JP-343441.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 DR WPI: 98-577037/49.
 PT A linked T cell epitope peptide - used for the treatment of
 PT sugi-pollinosis
 PS Claim 10; Page 5; 21pp; Japanese.
 CC W80339-58 represent epitopes for T cells, derived from the sugi allergen
 CC proteins Cryj1 (W80339-44, W80350-53 and W80356-58) and Cryj2 (W80345-49
 CC and W80354-55). The peptides are useful for the treatment of
 CC sugi-pollinosis, an allergic reaction of the body to pollen.
 SQ Sequence 47 AA;

Query Match 29.9%; Score 121.5; DB 1; Length 47;
 Best Local Similarity 58.7%; Pred. No. 1e-09; Indels 11; Gaps 2;
 Matches 27; Conservative 4; Mismatches 4; Indels 11; Gaps 2;

OY 34 FASKNFHLOKNTIGTGRIISLKLTSKGIIASRRVDGIIAYONPASP 79
 DB 13 FASKNFHLOKNT-----MKVT---VAFNFGGIIAYONPASP 47

RESULT 8

W80357 standard; peptide; 47 AA.
 ID W80357
 AC W80357:
 DT 11-JAN-1999 (first entry)
 DE Sugi allergen protein Cryj1 derived epitope for T cells.
 KW T cell epitope; sugi allergen proteins Cryj1; Cryj2; treatment;
 KM sugi-pollinosis; allergic reaction; pollen.
 OS Synthetic.
 PN J10259198-A.
 PD 29-SEP-1998.
 PF 22-DEC-1997: JP-343441.
 PR 24-DEC-1996: JP-343441.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 DR WPI: 98-577037/49.
 PT A linked T cell epitope peptide - used for the treatment of
 PT sugi-pollinosis
 PS Claim 10; Page 5; 21pp; Japanese.
 CC W80339-58 represent epitopes for T cells, derived from the sugi allergen
 CC proteins Cryj1 (W80339-44, W80350-53 and W80356-58) and Cryj2 (W80345-49
 CC and W80354-55). The peptides are useful for the treatment of
 CC sugi-pollinosis, an allergic reaction of the body to pollen.
 SQ Sequence 47 AA;

Query Match 29.6%; Score 120; DB 1; Length 47;
 Best Local Similarity 40.5%; Pred. No. 1.6e-09; Indels 44; Gaps 2;
 Matches 32; Conservative 2; Mismatches 1; Indels 44; Gaps 2;
 OY 1 MKVTAFNFGPNNRVRIFIKRVSNVLIHGRRIDFASKNFHLOKNTIGTGRIISLKLTSKG 60
 DB 1 MKVTAFNFGP-----FASKNFHLOKNT-----KLTSKG 17
 OY 61 IASRRVDGIIAYONPASP 79
 DB 18 IAS-CLNGLVHANNPASP 35

RESULT 9

W80356 standard; peptide; 47 AA.
 ID W80356
 AC W80356:
 DT 11-JAN-1999 (first entry)
 DE Sugi allergen protein Cryj1 derived epitope for T cells.
 KW T cell epitope; sugi allergen proteins Cryj1; Cryj2; treatment;
 KM sugi-pollinosis; allergic reaction; pollen.
 OS Synthetic.
 PN J10259198-A.
 PD 29-SEP-1998.
 PF 22-DEC-1997: JP-343441.
 PR 24-DEC-1996: JP-343441.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 DR WPI: 98-577037/49.
 PT A linked T cell epitope peptide - used for the treatment of
 PT sugi-pollinosis
 PS Claim 10; Page 5; 21pp; Japanese.
 CC W80339-58 represent epitopes for T cells, derived from the sugi allergen
 CC proteins Cryj1 (W80339-44, W80350-53 and W80356-58) and Cryj2 (W80345-49
 CC and W80354-55). The peptides are useful for the treatment of
 CC sugi-pollinosis, an allergic reaction of the body to pollen.
 SQ Sequence 47 AA;

Query Match
Best Local Similarity 28.1%; Score 114; DB 1; Length 47;
Matches 25; Conservative 0; Mismatches 0; Indels 22; Gaps 1;

QY 34 FASKNFHLQKNTIGTRISLKSGLASRRVGGTAAYONPASWK 80
DB 1 FASKNFHLQKNTIGTRISLKSGLASRRVGGTAAYONPASWK 25

RESULT 10

RS3690
ID RS3690 standard; Protein; 514 AA.
AC RS3690;
DT 01-FEB-1995 (first entry)
DE Japanese cedar pollen allergen Cry j II.
KW Cedar pollinosis; diagnostic.
OS Cryptomeria japonica.
PN WO9411512-A.
PD 26-MAY-1994.
PE 12-NOV-1993; U11000.
PR 12-NOV-1992; US-975179.
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
PI Brauer A, Kuo M, Pollock J, Yeung S;
DR WPI; 94-183513/22.
N-PSDB; 066048.
PT Allergenic Cry j II protein and fragments from Japanese cedar
pollen - used to diagnose, treat and prevent Japanese cedar
pollinosis
PS Claim 2; Fig 4; 89pp; English.
CC The sequence is of a Japanese cedar pollen allergen Cry j
II. The protein and its fragments can be used for diagnosis and
treatment of Japanese cedar pollinosis and to identify similar
sequences in other plants.
CC See also R53692-6.
SQ Sequence 514 AA;

Query Match
Best Local Similarity 24.4%; Score 99; DB 1; Length 514;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTG 49
DB 236 IDIFASKNFHLQKNTIGTG 254

RESULT 11

RS69792
ID R69792 standard; Protein; 514 AA.
AC R69792;
DT 27-SEP-1995 (first entry)
DE Japonicum allergen.
KW Japonicum allergen; induced histamine release; antiallergic peptide;
OS Japonicum sp.
PN WO9502412-A.
PD 26-JAN-1995.
PE 15-JUL-1994; J01164.
PR 16-JUL-1993; JP-177008.
PR 01-SEP-1993; JP-217725.
PR 07-APR-1994; JP-069336.
PA (MEIP) MEIJI MILK PROD CO LTD.
PI Kuno K, Kohno Y, Komiyama N, Sone T;
DR WPI; 95-067159/09.
N-PSDB; 084045, 084046.
PT Peptide antiallergic agent - inhibits cross-linking of allergen
with IgE antibody
Example 3; Pages 27-28; 46pp; Japanese.
CC 084045 encodes R69792 Japonicum allergen, from which the antiallergic
peptides R69845-R69809 were derived. The peptides ability to inhibit
the cross-linking of an allergen, to an IgE antibody can be used in
the prevention and treatment of allergic diseases.

SQ Sequence 514 AA;

Query Match
Best Local Similarity 24.4%; Score 99; DB 1; Length 514;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTG 49
DB 236 IDIFASKNFHLQKNTIGTG 254

RESULT 12

RS69791
ID R69791 standard; Protein; 460 AA.
AC R69791;
DT 27-SEP-1995 (first entry)
DE Japonicum allergen residues 55-514.
KW Japonicum allergen; residues 55-514; induced histamine release;
antiallergic peptide; IgE cross-linking inhibition.
OS Japonicum sp.
PN WO9502412-A.
PD 26-JAN-1995.
PE 15-JUL-1994; J01164.
PR 16-JUL-1993; JP-177008.
PR 01-SEP-1993; JP-217725.
PR 07-APR-1994; JP-069336.
PA (MEIP) MEIJI MILK PROD CO LTD.
PI Kuno K, Kohno Y, Komiyama N, Sone T;
DR WPI; 95-067159/09.
N-PSDB; 084044.
PT Peptide antiallergic agent - inhibits cross-linking of allergen
with IgE antibody
PS Disclosure; Pages 26-27; 46pp; Japanese.
CC 084044 encodes R69791 Japonicum allergen residues 55-514, from
which the antiallergic peptides R69845-R69809 were derived.
CC The peptides ability to inhibit the cross-linking of an allergen,
to an IgE antibody can be used in the prevention and treatment of
allergic diseases.
SQ Sequence 460 AA;

Query Match
Best Local Similarity 24.4%; Score 99; DB 1; Length 460;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTG 49
DB 182 IDIFASKNFHLQKNTIGTG 200

RESULT 13

RS74333
ID R74333 standard; Protein; 514 AA.
AC R74333;
DT 01-NOV-1995 (first entry)
DE Japanese cedar pollen allergen.
KW Japanese cedar pollen allergen;
OS Cryptomeria japonica.
PN EP-65500-A.
PD 03-NOV-1994; 308117.
PR 05-NOV-1993; JP-299151.
PR 20-DEC-1993; JP-344596.
PR 27-DEC-1993; JP-346814.
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
PI Kurimoto M, Namba M, Torigoe K;
DR WPI; 95-195588/26.
N-PSDB; 090156.
PT New Japanese cedar pollen allergen polypeptide - and DNA coding for it,
useful for treatment and diagnosis of cedar pollen allergy
Claim 5; Page 26-28; 41pp; English.

CC The gene encoding an allergen of Japanese cedar pollen was isolated
 CC by PCR amplification using primers based on portions of the allergen
 CC protein. The gene was used for recombinant allergen production in
 CC E. coli (vector plasmid pK-223-3).
 SQ Sequence 514 AA:

Query Match 24.4%; Score 99; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 2.5e-05;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 IDIFASKNFHLOKNTIGTG 49
 ||||||||||||||||
 DB 236 IDIFASKNFHLOKNTIGTG 254

RESULT 14

R81586
 ID R81586 standard; Protein; 514 AA.
 AC R81586;
 UA 24-MAY-1996 (first entry)
 DE Cedar pollen allergen; immunoglobulin E; IgE; T-cell epitope;
 KW Cedar pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
 OS Cryptomeria japonica.
 PN EP-700929-A2.
 PD 13-MAR-1996.
 PR 08-SEP-1995; 306295.
 PR 10-SEP-1994; JP-242137.
 PR 14-JUL-1995; JP-200221.
 PR 14-JUL-1995; JP-200204.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PI Hino K, Saito S, Taniguchi Y;
 PI WPI: 96-140976/15.
 DR New peptide(s) derived from cedar pollen allergens - activate
 PR allergen-specific T-cells, but not allergen-specific IgE antibodies,
 PT used for treating cedar pollinosis.
 PS Claim 5; Page 29-30; 36pp; English.
 CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (R81586) and B (R81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 peptides (R81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (R81573-79) essential
 CC for T-cell recognition, and homologous peptides (R81588-96) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.
 SQ Sequence 514 AA:

Query Match 24.4%; Score 99; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 2.5e-05;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 IDIFASKNFHLOKNTIGTG 49
 ||||||||||||||||
 DB 236 IDIFASKNFHLOKNTIGTG 254

RESULT 15

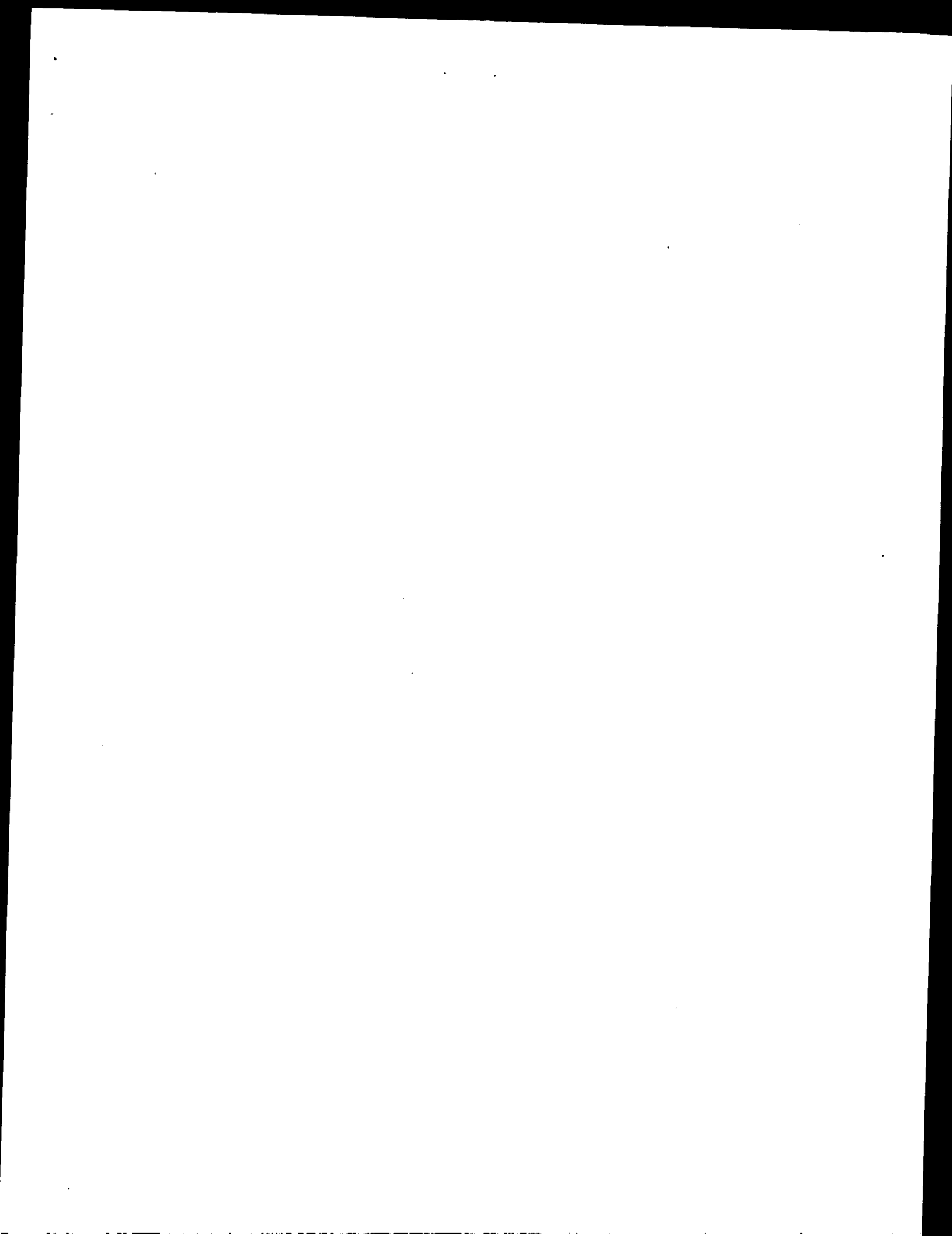
R93599
 ID R93599 standard; Protein; 514 AA.
 AC R93599;
 UA 16-AUG-1996 (first entry)
 DE Japan cedar pollen Cry j II allergen.
 KW Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;
 OS Sugi pollinosis; diagnosis; treatment.
 OS Cryptomeria japonica.
 FH Key location/Qualifiers
 FT peptide 1..54
 FT /*label- sig_peptide
 FT 55..514
 FT /*label- mat_protein
 PN J08047392-A.

PD 20-FEB-1996.
 PF 07-NOV-1994; 297840.
 PR 05-NOV-1993; JP-276773.
 PR 26-MAY-1994; JP-134868.
 PA (MEIP) MEIJU MILK PROD CO LTD.
 DR WPI: 96-166249/17.
 DR N-PSDB; T18102.
 PT Japan cedar pollen allergen Cry j II epitope - comprises at least
 PR part of specified 460 amino acid protein
 PS Claim 1; Page 10-11; 17pp; Japanese.
 CC R93599 is a Japan cedar pollen Cry j II allergen which is useful
 CC in the diagnosis, prevention and treatment of Sugi pollinosis,
 CC the allergic reaction to Japan cedar pollen. Significant regions of
 CC the allergen were identified using overlapping peptides of the full
 CC epitope derived from a Cry j II antigen-specific T cell line
 CC (see R97871-R97960). Amino acids 66-80 (R97884) and 186-200 (R978908)
 CC of the full mature 460 amino acid allergen are the most allergenic
 CC of the 90 peptides tested.
 SQ Sequence 514 AA:

Query Match 24.4%; Score 99; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 2.5e-05;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 IDIFASKNFHLOKNTIGTG 49
 ||||||||||||||||
 DB 236 IDIFASKNFHLOKNTIGTG 254

Search completed: September 29, 1999, 13:35:10
 Job time: 2050 sec



GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 1999, 13:01:00 ; Search time 31.02 Seconds
(without alignments)
103.328 Million cell updates/sec

Title: US-09-142-524-1

Perfect score: 406
Sequence: 1 MKVYAFNQFGNRRVFIKR.....IASRRVDGIATAYQNPASMK 80

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database: PIR.60:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	24.4	514	2 S48730	Cry j II protein -
2	99	24.4	514	2 JC2498	second major aller
3	71	17.5	374	2 JC2124	major allergen Cry
4	71	17.5	374	2 JC2123	major allergen Cry
5	70	17.2	152	1 R3KR18	ribosomal protein
6	70	17.2	152	2 S30393	ribosomal protein
7	70	17.2	152	2 I76666	ribosomal protein
8	66.5	16.4	196	2 S54580	probable membrane
9	66.5	16.4	350	2 S57065	sensory transducti
10	65	16.0	498	2 J02353	glycoprotein E pre
11	64	15.8	222	2 S78174	probable transport
12	62.5	15.4	1280	2 A36466	1-phosphatidylinos
13	62.5	15.4	434	2 S29612	pectate lyase - tr
14	62	15.3	173	1 J0ECP6	fimbrial protein p
15	61.5	15.1	633	2 S62737	topoisomerase I -
16	61	15.0	173	2 B27743	pape fimbrial prot
17	60.5	14.9	110	2 D48562	hypothetical nucle
18	60	14.8	627	1 TYMSVY	transforming prote
19	60	14.8	725	2 T00484	hypothetical prote
20	60	14.8	725	2 S60711	band-6-protein - h
21	60	14.8	295	2 S60711	band-6-protein - b
22	59.5	14.7	841	2 S34624	glycogen phosphory
23	59.5	14.7	254	2 D63945	ABC transporter, A
24	59	14.5	839	2 S53913	FKB2 protein - yea
25	58.5	14.4	1291	2 TYHUVV	transforming prote
26	58.5	14.4	478	2 S00666	1-phosphatidylinos
27	58.5	14.4	478	2 S61986	subtilisin-like pr
28	58.5	14.4	380	2 S61986	CMP-NeuAc:GM3 alph
29	58	14.3	565	2 J0C321	hemagglutinin prec
30	58	14.3	152	1 HMIWDE	ribosomal protein
31	58	14.3	578	2 S37496	ABC transporter, A
32	58	14.3	1396	2 F71844	hypothetical prote
33	58	14.3	1396	2 G70598	xanthine dehydroge
34	58	14.3	1396	2 S73093	hypothetical prote
35	57.5	14.2	1070	2 S46755	1-phosphatidylinos
36	57.5	14.2	1070	2 A31317	hypothetical prote
37	57	14.0	103	2 E70170	UTP--glucose-1-pho
38	57	14.0	511	1 XND0U	dihydrolipoamide S
39	57	14.0	314	2 J0E0127	catechol 2,3-dioxy

40 57 14.0 616 2 JC4084
41 57 14.0 379 2 S55845
42 56.5 13.9 167 2 B23093
43 56.5 13.9 994 2 A47474
44 56.5 13.9 1163 2 S07137
45 56.5 13.9 215 2 A48463

prolyl endopeptidase
3-isopropylmalate
glycogen phosphory
NAD+ ADP-ribosyltr
DNA-directed RNA p
Ras-like GTP-bind1

ALIGNMENTS

RESULT 1
S48730
Cry j II protein - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 21-Jul-1995
C:Accession: S48730
R:Namba, M.; Kurose, M.; Torigoe, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.;
FEBS Lett. 353, 124-128, 1994
A:Title: Molecular cloning of the second major allergen, Cry j II, from Japanese ceda
A:Reference number: S48730; MUID:95010777
A:Accession: S48730
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-514 <NAM>

Query Match 24.4%; Score 99; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLOKNTIGTG 49
DB 236 IDIFASKNFHLOKNTIGTG 254

RESULT 2
JC2498
second major allergen Cry j II precursor - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 12-Sep-1997
C:Accession: JC2498; PC2346; A60147
R:Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994
A:Title: cDNA cloning and expression of Cry j II, the second major allergen of Japan
A:Reference number: JC2498; MUID:94271186
A:Accession: JC2498
A:Molecule type: mRNA
A:Residues: 1-514 <KOK>
A:Cross-references: DDBJ:D29772; NID:g506857; PID:d1006734; PID:g506858
A:Accession: PC2346
A:Molecule type: protein
A:Residues: 52-61 <KO2>
R:Sakaguchi, M.; Inouye, S.; Tanai, M.; Ando, S.; Usui, M.; Matuhasi, T.
Allergy 45, 309-312, 1990
A:Title: Identification of the second major allergen of Japanese cedar pollen.
A:Reference number: A60147; MUID:90342988
A:Accession: A60147
A:Molecule type: protein
A:Residues: 55-64 <SAK>
C:Keywords: glycoprotein; pollen
F:1-54/Domain: signal sequence #status predicted <SIG>
F:55-460/Product: second major allergen Cry j #status predicted <MAY>
F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.4%; Score 99; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLOKNTIGTG 49
DB 236 IDIFASKNFHLOKNTIGTG 254

A: Variety: PCC 6803 #sequence_revision 25-Apr-1997 #text_change 21-Aug-1998
C: Date: 25-Apr-1997
C: Accession: S75065
R: Kaneko, T.: Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, R.; Kaneo, T.; Sato, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu, O.; K. ; Okumura, S.;
DNA Res. 3, 109-136, 1996
A: Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
S: Reference number: S74322; MUID: 97061201
A: Accession: S75065
A: Status: nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-350 <KAN>
A: Cross-references: EMBL: D9910; GB: AB001339; NID: g1652956; PID: d1018660; PID: g165301
A: Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C: GenBank:
A: Start codon: GTG

```

Query Match          16.3%; Score 66; DB 2; Length 350;
Best Local Similarity 39.5%; Pred. No. 2.4; Indels 4; Gaps 2;
Matches 17; Conservative

QY      14  RRVFIKRVSNVLIH--GRRIDI-FASKNFHLQKNTIGTGRI 52
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db      246  QRVFNMLLTALNHSPRGKRVISLTSKNHFOVQLVDEGRGI 288

RESULT 10
JQ2353
glycoprotein E precursor - turkey herpesvirus
N.Alternate names: ORF 8 protein
C.Species: turkey herpesvirus
C.Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 09-Sep-1994
C.Accession: JQ2353
C.Creation: JQ2353
C.Revision: V.; Darteil, R.; Audonnet, J.C.; Smith, G.D.; Riviere, M.; Pastorek, J.;
J. Gen. Virol. 74, 2151-2162, 1993
A.Title: The complete sequence and gene organization of the short unique region of
A.Reference number: JQ2346
A.Accession: JQ2353
A.Molecule type: DNA
A.Residues: 1-498 <ZEL>
C.Keywords: glycoprotein; transmembrane protein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-498/Product: glycoprotein E #status predicted <MAN>
F:386-412/Domain: transmembrane #status predicted <TM>
F:44,60,133,148,370/Binding site: carbohydrate (Asn) (covalent) #status predicted

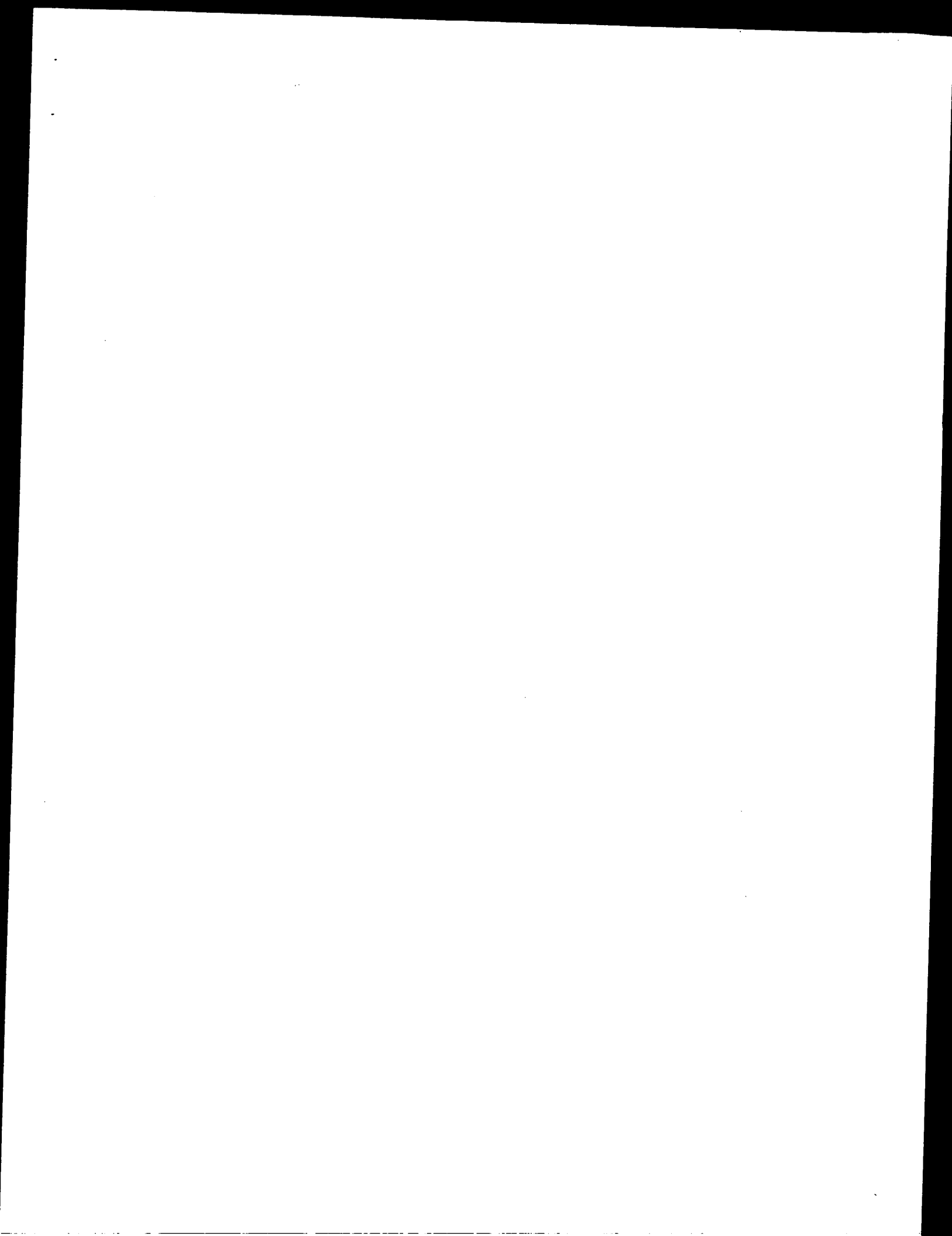
Query Match          16.0%; Score 65; DB 2; Length 498;
Best Local Similarity 31.6%; Pred. No. 4.7;
Matches 18; Conservative 10; Mismatches 25; Indels 4; Gaps 2;

QY      18  IKRVSNIHIGRRIDIFASKNFHLQKNT-IGTGRRSIKLTSGKISRVDGIILAY 73
      | | | :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db      92  ILRVDFVNVNCSGLDRIASSOFHYIPNVIIIGTGRKELTIFN--ATSOIAGVYTRY 145

RESULT 11
S78174
probable transport protein yeyW - Reclinomonas americana (ATCC 50394) mitochondrion
C.Species: mitochondrion Reclinomonas americana
A.Variety: ATCC 50394
C.Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 17-Mar-1999
C.Accession: S78174
C.Creation: S78174
C.Revision: B.F.; Burger, G.; O'Kelly, C.J.; Cedergren, R.; Golding, G.B.; Lemieux, C
Nature 387, 493-497, 1997
A.Title: An ancestral mitochondrial DNA resembling a eubacterial genome in miniatu
A.Reference number: S78127; MUID:97311393
A.Accession: S78174
A.Status: nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Residues: 1-222 <LAN>

```

C;Gene: pape; prse
A;Gene: pape; prse
C;Superfamily: pape_fimbrial_protein
C;Keywords: fimbria; pili adhesion
F;1-24/Domain: signal sequence #status predicted <Sig>



GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 1999, 13:14:01 ; Search time 21.98 Seconds

(without alignments)
102.887 Million cell updates/sec

Title: US-09-142-524-1

Perfect score: 406

Sequence: 1 MVTVAFNFGPFRVFIK.....IASRVDGLIAYQNPASMK 80

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database: SwissProt_37.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	24.4	514	1	MPA2_CRYJA
2	71	17.5	374	1	SBP_CRYJA
3	70	17.2	132	1	RS18_HUMAN
4	66.5	16.4	156	1	YMO7_YEAST
5	62.5	15.4	434	1	PEL_LILLO
6	62.5	15.4	1290	1	PIP4_HUMAN
7	62.5	15.4	145	1	YPI2_AGRTE
8	62	15.3	173	1	PAPE_ECOLI
9	61.3	15.1	633	1	TOP1_THEMA
10	61	15.0	152	1	RS18_DROME
11	60.5	14.9	110	1	Y12K_SMSV4
12	59.5	14.7	842	1	PHS2_RAT
13	59	14.5	862	1	EKH2_YEAST
14	58.5	14.4	1281	1	PIP4_YEAST
15	58.5	14.4	478	1	YSP3_YEAST
16	58	14.3	665	1	HEMA_LADE1
17	58	14.3	646	1	NODD_RHIS3
18	58	14.3	152	1	RS18_ARATH
19	58	14.3	1070	1	YHVA_YEAST
20	57.5	14.2	207	1	PIP4_MOUSE
21	57.5	14.2	1290	1	PIP4_RAT
22	57	14.0	567	1	ODP2_HAEIN
23	57	14.0	511	1	UDPG_DICDI
24	57	14.0	379	1	YLED_DBOC
25	56.5	13.9	339	1	BMP4_BORBU
26	56.5	13.9	342	1	MEMB_METTR
27	56.5	13.9	842	1	PHS2_HUMAN
28	56.5	13.9	994	1	PPOL_DROME
29	56.5	13.9	215	1	RAN_BRUMA
30	56.5	13.9	1163	1	RPOD_PEA
31	56	13.8	428	1	AROA_BACSU
32	56	13.8	486	1	LACI_BOTCI
33	56	13.8	633	1	NODD_RHISB
34	56	13.8	843	1	PHS3_HUMAN
35	56	13.8	267	1	RS3_MYCGE
36	55.5	13.7	842	1	PHS2_RABIT
37	55.5	13.7	1305	1	RRLP_AHSV9
38	55.5	13.7	277	1	SET_HUMAN
39	55	13.5	185	1	NIR_LEPWC
40	55	13.5	629	1	ODP2_ECOLI
41	55	13.5	352	1	RECA_CHLTR
42	54.5	13.4	470	1	DLDI_BACSU
43	54.5	13.4	413	1	L756_CAEEL

44 54.5 13.4 532 1 PRCK_AKASU 009460 anaerobiosp
45 54 13.3 1296 1 ASAL_ENTFA P17953 enterococcu

ALIGNMENTS

```

RESULT 1
MPA2_CRYJA STANDARD; PRT; 514 AA.
AC P43212;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE POSSIBLE POLYGALACTURONASE PRECURSOR (BC 3.2.1.15) (PG) (PECTINASE)
DE (MAJOR POLLEN ALLERGEN CRY J 2) (CRY J II).
OS CRYPTOMERIA JAPONICA (JAPANESE CEDAR).
OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTES; EMBRYOPHYTES; TRACHEOPHYTES;
OC EUPHYLOPHYTES; SPERMATOPHYTES; CONIFEROPSIDA; CONIFERALES;
OC TAXODIACEAE; CRYPTOMERIA.
RN
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE-POLLEN.
RX MEDLINE: 95010777.
RA NAKABA M., KUROSE M., TORIGOE K., HINO K., TANIGUCHI Y., FUKUDA S.,
RA USUI M., KURIMOTO M.;
RT "Molecular cloning of the second major allergen, Cry j II, from
RT Japanese cedar pollen."
RL FEBS LETT. 353:124-128(1994).
RN
RP SEQUENCE FROM N.A.
RP TISSUE-POLLEN.
RX MEDLINE: 94271166.
RA KOMIYAMA N., SONE T., SHIMIZU K., MORIKUBO K., KINO K.;
RT "CDNA cloning and expression of Cry j II the second major allergen of
RT Japanese cedar pollen."
RL BIOCHEM. BIOPHYS. RES. COMMUN. 201:1021-1028(1994).
RN
RP SEQUENCE OF 55-64.
RX MEDLINE: 90342988.
RA SAKAGUCHI M., INOUE S., TANAI M., ANDO S., USUI M., MATUNASI T.;
RT "Identification of the second major allergen of Japanese cedar
RT pollen."
RL ALLERGY 45:309-312(1990).
CC -1- CATALYTIC ACTIVITY: RANDOM HYDROLYSIS OF 1,4-ALPHA-D-
CC SUBCELLULAR LOCATION: IN PECTATE AND OTHER GALACTURONANS.
CC -1- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
CC (POLYGALACTURONASES).
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D37765; G577696; -
DR EMBL: D29772; G506858; -
DR PROSITE: PS00502; POLYGALACTURONASE; 1.
DR PFAM: PF00295; Pg; 1.
KW HYDROLASE; GLYCOSIDASE; CELL WALL; SIGNAL; ZYMOGEN; FRUIT RIPENING;
KW AMYLOPLAST; GLYCOPROTEIN; ALLERGEN.
FT SIGNAL 1 45
FT PROPEP 2 45
FT CHAIN 46 433
FT PROPEP 434 514
FT ACT_SITE 278 278
FT CARBOHYD 460 460
FT CARBOHYD 472 472
FT CONFLICT 5 5
FT CONFLICT 12 12
FT V -> L (IN REF. 2).
FT V -> L (IN REF. 2).

```

```

CC CONFLICT 34 35 DI -> VV (IN REF. 2).
CC CONFLICT 37 37 Q -> K (IN REF. 2).
CC CONFLICT 88 88 K -> N (IN REF. 2).
CC CONFLICT 98 98 N -> S (IN REF. 2).
CC CONFLICT 451 451 K -> E (IN REF. 2).
CC CONFLICT 454 454 G -> R (IN REF. 2).
CC CONFLICT 504 504 M -> I (IN REF. 2).
CC CONFLICT 507 507 R -> C (IN REF. 2).
CC SEQUENCE 514 AA: 56645 MW; 467FE33 CRC32;

Query Match
Best Local Similarity 100.0%; Score 99; DB 1; Length 514;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 31 IDIFASKNFHLOKNTIGTC 49
236 IDIFASKNFHLOKNTIGTC 254

RESULT 2
SBP_CRYJA STANDARD; PRT; 374 AA.
ID SBP_CRYJA
AC P18632;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 15-DEC-1996 (REL. 37, LAST ANNOTATION UPDATE)
DE SUGI BASIC PROTEIN PRECURSOR (SBP) (MAJOR ALLERGEN CRY J 1) (CRY J 1).
OS CRYPTOMERIA JAPONICA (JAPANESE CEDAR).
OC EUKARYOTA: VIRIDIPHYTES: STREPTOPHYTA: TRACHEOPHYTA:
OC EUPHYLOPHYTES: SPERMATOPHYTA: CONIFEROPSIDA: CONIFERALES;
OC TAXODIACEAE: CRYPTOMERIA.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=POLLEN;
RX MEDLINE: 94183234.
RA SONE T., KOMITAMA N., SHIMIZU K., KUSAKABE T., MORIKUBO K.,
RA KINO K.,
RA "Cloning and sequencing of cDNA coding for Cry j 1, a major allergen
RT of Japanese cedar pollen."
RL BIOCHEM. BIOPHYS. RES. COMMUN. 199:619-625(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=POLLEN;
RA NAKAMA M., KUROSE M., TORIGOE K., FUKUDA S., KURIMOTO M.;
RL SUBMITTED (JUL-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE OF 22-41.
RC TISSUE=POLLEN;
RX MEDLINE: 89031257.
RA TANAI M., ANDO S., USUI M., KURIMOTO M., SAKAGUCHI M., INOUE S.,
RA MATOHASHI T.;
RA "N-terminal amino acid sequence of a major allergen of Japanese cedar
RT pollen (Cry j 1)."
RL FEBS LETT. 239:329-332(1988).
RN [4]
RP CARBOHYDRATES.
RX MEDLINE: 95003748.
RA HIRAKATA A., MATSUMOTO I., KOJIMA K., OGAWA H.;
RT "Antigenicity of the oligosaccharide moiety of the Japanese cedar
RL INT. ARCH. ALLERGY IMMUNOL. 105:198-202(1994).
CC -1- PTM: CONTAINS FUCOSE/XILOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.
CC -1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR
CC POLLEN. THE MOST COMMON POLLEN ALLERGEN IN JAPAN.
CC -1- THE SEQUENCE OF CRY J 1 FORM A IS SHOWN HERE. FORM B DIFFERS IN
CC SIX POSITIONS.
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A I/AMB A II/CRY J I SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: D26544; G493632; -
CC EMBL: D26545; G493636; -
CC EMBL: D34639; G516726; -
CC PIR: A44773; A44773.
CC DR PFAM: PF00544; pec_lyase; 1.
CC DR ALLERGEN; GLYCOPROTEIN; MULTIGENE FAMILY; SIGNAL.
CC SIGNAL 1 21
CC CHAIN 22 374 SUGI BASIC PROTEIN.
CC VARIANT 12 12 L -> F (IN CRY J 1-B).
CC VARIANT 143 143 H -> Y (IN CRY J 1-B).
CC VARIANT 202 202 S -> T (IN CRY J 1-B).
CC VARIANT 221 221 L -> S (IN CRY J 1-B).
CC VARIANT 358 358 Q -> Q (IN CRY J 1-B).
CC VARIANT 361 361 K -> Q (IN CRY J 1-B).
CC CARBOHYD 158 158 POTENTIAL.
CC CARBOHYD 191 191 POTENTIAL.
CC CARBOHYD 293 293 POTENTIAL.
CC SEQUENCE 374 AA: 40645 MW; DC088629 CRC32;

Query Match
Best Local Similarity 17.5%; Score 71; DB 1; Length 374;
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Db 1 MKYVAFNFGPRFRVIRVSVIHH 27
234 MKYVAFNFGPRGCGMPRARKGLVH 260

RESULT 3
RS18 HUMAN STANDARD; PRT; 152 AA.
ID RS18 HUMAN
AC P25232;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE 40S RIBOSOMAL PROTEIN S18 (KE-3) (KE3).
GN RPS18.
OS HOMO SAPIENS (HUMAN). RATTUS NORVEGICUS (RAT), MUS MUSCULUS (MOUSE),
OS AND SUS SCROFA (PIG).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
OC PRIMATES: CATARRHINI: HOMINIDAE: HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=HUMAN; TISSUE=PLACENTA;
RX MEDLINE: 93181276.
RA CHASSIN D., BELLET D., ROMAN A.;
RT "The human homolog of ribosomal protein S18."
RL NUCLEIC ACIDS RES. 21:745-745(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=RAT; STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX MEDLINE: 91337062.
RA CHAN Y.-L., PAZ V., WOOL I.G.;
RT "The primary structure of rat ribosomal protein S18."
RL BIOCHEM. BIOPHYS. RES. COMMUN. 178:1212-1218(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=MOUSE;
RX MEDLINE: 92182530.
RA MACMURRAY A.J., SHIN H.S.;
RT "The murine mhc encodes a mammalian homolog of bacterial ribosomal
RL protein S18."
RN MAMM. GENOME 2:87-95(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=PIG;
RA KIMURA M., KANAKAMI K., SUZUKI H., HAMASIMA N.;

```


Best Local Similarity 28.4%; Pred. No. 4.2;
Matches 25; Conservative 11; Mismatches 33; Indels 19; Gaps 4;

QY 1 MKVYAFNPGPN-----RRVFKRVSNIHGRIDIFASKNFHLQKTTGTGR- 52
DB 296 MOVTAFNHFGRVGRMPRCRGFEVHYVNDYTHWIMYVAGSOH---PILISQGNRY 351
QY 52 ISLKTSGKIASRRVDGIIAAYONPASW 79
DB 352 IAPHEAKEVTKR-----DYAEPAEW 373

RESULT 6

PIPA_HUMAN STANDARD; PRT; 1290 AA.
AC P19174;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE 1-PHOSPHATIDYLINOSITOL-4,5-BISPHOSPHATE PHOSPHODIESTERASE GAMMA 1
DE (EC 3.1.4.11) (PLC-GAMMA-1) (PHOSPHOLIPASE C-GAMMA-1) (PLC-II)
DE (PLC-148).
GN PLCG1 OR PLC1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
AN [1]
AA SEQUENCE FROM N.A.
AU ISSUE-VEIN, AND BRAIN;
RX BURGESS W.H., DIONNE C.A., KAPLOW J.M., MUDD R., FRIESEL R.,
PA ZILBERSTEIN A., SCHLESINGER J., JAYE M.;
RA "Characterization and cDNA cloning of phospholipase C-gamma, a major
RT substrate for heparin-binding growth factor 1 (acidic fibroblast
RT growth factor)-activated tyrosine kinase.";
RL MOL. CELL. BIOL. 10:4770-4777(1990).
RN [2]
RA BRIDGEMAN A.;
RT SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RX STRUCTURE BY NMR OF SH3 DOMAIN.
RX MEDLINE; 93208890.
RA KOHDA D., HATANAKA H., ODAKA M., MANDIYAN V., ULLRICH A.,
RA SCHLESINGER J., INAGAKI F.;
RT "Solution structure of the SH3 domain of phospholipase C-gamma.";
RL CELL 72:953-960(1993).
CC -1- FUNCTION: PLC-GAMMA IS A MAJOR SUBSTRATE FOR HEPARIN-BINDING
CC GROWTH FACTOR 1 (ACIDIC FIBROBLAST GROWTH FACTOR)-ACTIVATED
CC TYROSINE KINASE.
CC -1- CATALYTIC ACTIVITY: 1-PHOSPHATIDYL-D-MYO-INOSITOL 4,5-BISPHOSPHATE
CC + H(2)O -> D-MYO-INOSITOL 1,4,5-TRIPHOSPHATE + DIACYLGLYCEROL.
CC -1- COFACTOR: REQUIRES CALCIUM.
CC -1- PTM: THE RECEPTOR-MEDIATED ACTIVATION OF PLC-GAMMA 1 AND 2
CC INVOLVES THEIR PHOSPHORYLATION BY TYROSINE KINASES IN RESPONSE
CC TO LIGATION OF A VARIETY OF GROWTH FACTOR RECEPTORS AND IMMUNE
CC SYSTEM RECEPTORS.
CC -1- THERE ARE AT LEAST SIX FORMS OF PLC ENZYMES.
CC -1- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN DIFFERENT FORMS
CC OF PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 PH DOMAINS. THE SECOND ONE IS IN TWO
CC PARTS.
CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M34667; G190038; -;
DR EMBL; AL022394; E1312192; -;
DR PIR; A36466; A36466;
DR PDB; 2HSP; 31-AUG-94.
DR PDB; 1HSQ; 31-AUG-94.
DR PDB; 17420; -;
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS50001; SH2; 2.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS50003; PH_DOMAIN; 2.
DR PROSITE; PS50004; C2_DOMAIN; 2; 1.
DR PROSITE; PS50007; PIP2C_X_DOMAIN; 1.
DR PROSITE; PS50008; PIP2C_Y_DOMAIN; 1.
DR PFAM; PF00017; SH2; 2.
DR PFAM; PF00018; SH3; 1.
DR PFAM; PF00168; C2; 1.
DR PFAM; PF00169; PH; 2.
DR PFAM; PF00387; PI-PLC-Y; 1.
DR PFAM; PF00388; PI-PLC-X; 1.
KW HYDROLASE; LIPID DEGRADATION; TRANSUGER; SH2 DOMAIN; SH3 DOMAIN;
KW CALCIUM-BINDING; PHOSPHORYLATION; 3D-STRUCTURE.
FT DOMAIN; 27 142
FT CA_BIND 165 176
FT DOMAIN 320 464
FT DOMAIN 489 523
FT DOMAIN 550 657
FT DOMAIN 668 756
FT DOMAIN 791 851
FT DOMAIN 895 931
FT DOMAIN 953 1070
FT DOMAIN 1075 1177
FT ACT_SITE 335 335
FT ACT_SITE 380 380
FT MOD_RES 771 771
FT MOD_RES 783 783
FT MOD_RES 1253 1253
SQ SEQUENCE 1290 AA; 148531 MW; D3DC5B9D CRC32;

Query Match

Best Local Similarity 28.1%; Score 62.5; DB 1; Length 1290;

Matches 18; Conservative 9; Mismatches 34; Indels 3; Gaps 1;

QY 5 VAFNPGPNRRVFKRVSNIHGRIDIFASKNFHLQ--KNITGTGRISLKTSGKI 61
DB 883 IAIPEGKNNRLFFVFSISMASVAMSLDVADSOELEDWVKIRREVAQTADARLTGKI 942
QY 62 ASRR 65
DB 943 MERR 946

RESULT 7

YPI2_AGR6 STANDARD; PRT; 145 AA.
ID YPI2_AGR6
AC Q44433;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHEICAL 15.8 KD PROTEIN IN PIN2 3'REGION (ORF2).
DE AGROBACTERIUM TUMEFACIENS.
OS PLASMID P19A6.
OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RHIZOBIACEAE GROUP;
OC RHIZOBIACEAE; AGROBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89213933.
RA KANDMOTO R.H., POWELL A.T., AKIYOSHI D.E., REGIER D.A.,
RA KENSTETTER R.A., NESTER E.W., HAWES M.C., GORDON M.P.;
RT "Nucleotide sequence and analysis of the plant-inducible locus pinf
RT from Agrobacterium tumefaciens.";
RL J. BACTERIOL. 171:2506-2512(1989).

```
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19352; G1088264; -
RW HYPOTHETICAL PROTEIN; PLASMID.
SQ SEQUENCE 145 AA; 15829 MW; 0BF23123 CRC32;
Oy Query Match 15.4%; Score 62.5; DB 1; Length 145;
Best Local Similarity 33.3%; Pred. No. 1.2;
Matches 17; Conservative 11; Mismatches 18; Indels 5; Gaps 2;
Db 1 IKRVSNVTHH-CRRIDFASKNFHQNKNTIGCRISIKLTSGKIASRVD 67
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1 MKRISTILVGVFPLATPYAADNTH-----TLGLSELEALTAGKPVYNTVD 47
RESULT 8
PAPE_ECOLI STANDARD: PRT: 173 AA.
AC P08407;
AD 01-AUG-1968 (REL. 08, CREATED)
DT 01-AUG-1968 (REL. 08, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE FIMBRIAL PROTEIN PAPE PRECURSOR.
GN PAPE.
GC ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERiaceae;
OC ESCHERICHIA.
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN-J96.
RX MEDLINE; 86149403.
RA LINDBERG F., LUND B., NORMARK S.;
RT "Gene products specifying adhesion of uropathogenic Escherichia coli
RT are minor components of pili.";
RL PROC. NATL. ACAD. SCI. U.S.A. 83:1891-1895(1986).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN-J96.
RX MEDLINE; 88169520.
RA LUND B., LINDBERG F., NORMARK S.;
RT "Structure and antigenic properties of the tip-located P pilus
RT proteins of uropathogenic Escherichia coli.";
RL J. BACTERIOL. 170:1887-1894(1988).
[3]
RN RP SEQUENCE FROM N.A.
RC STRAIN-J96.
RX MEDLINE; 93023852.
RA MARKLUND B.-I., TENNENT J.M., GARCIA E., HAMERS A., BAGHA M.,
RA LINDBERG F., GAASTRA W., NORMARK S.;
RT "Horizontal gene transfer of the Escherichia coli pap and prs pili
RT operates as a mechanism for the development of tissue-specific
RT adhesive properties.";
RL MOL. MICROBIOL. 6:2225-2242(1992).
RL CC -1- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING
CC FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5
CC MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO
CC COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE MINOR COMPONENTS OF PILI.
CC PILI WITH A DEFECTIVE PAPE GENE WILL HAVE LOW ADHESIVE CAPACITY
CC OR NONE; HOWEVER, THE BINDING PROPERTY OF THE WHOLE CELL WILL NOT
CC BE AFFECTED.
CC -1- SUBCELLULAR LOCATION: SECRETED; TIP OF THE PILI.
CC -1- DISEASE: STRAINS OF E. COLI THAT CAUSE INFECTION OF THE HUMAN
CC URINARY TRACT PRODUCE PAP-PILI WHICH ARE HAIR-LIKE APPENDAGES
CC CONSISTING OF ABOUT 1000 HELICALLY ARRANGED SUBUNTS OF THE
```

```
CC PROTEIN PAPE. THESE PILI MEDIATE BINDING TO DIGITALTOSIDE-
CC CONTAINING GLYCOPOLIPIDS PRESENT ON THE EPITHELIAL CELLS WHICH
CC LINE THE URINARY TRACT.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X61238; G42527; -.
CC DR EMBL; M13239; G147075; -.
CC DR EMBL; M20146; G147090; -.
CC DR EMBL; X61239; G42299; -.
CC DR PIR; A25134; YOECEP.
CC DR PFAM; PF00419; Fimbrin1; 1.
CC KW FIMBRIN; CELL ADHESION; SIGNAL.
CC FT SIGNAL
CC FT CHAIN 1 24 POTENTIAL.
CC SQ SEQUENCE 173 AA; 18569 MW; DA6F5994 CRC32.
CC -----
Query Match 15.3%; Score 62; DB 1; Length 173;
Match Local Similarity 27.0%; Pred.No. 1.7; Indels 8; Gaps 3;
Matches 20; Conservative 11; Mismatches 35;
QY 1 MKVTV-AFNQFGNRRVFVKRVSNTIIGHRIDIFASKNFHLQKNTGTGRISLTSG 59
Db 79 MKVTIATNTNY--NNALIVQNTSNTSSDGLLYVLNS-----NAGNIGTAITGTFPTTG 131
QY 60 KIASRVVDGIITAY 73
Db 132 KITGNNAKDITSLH 145
DE DE DNA TOPoisOMERASE I (EC 5.99.1.2) (OMEGA-PROTEIN) (RELAXING ENZYME)
DE (UNWINDING ENZYME) (SMIVELASE).
GN TOPA.
OS THERMOTOGA MARTITA.
OC BACTERIA; THERMOTOGALES; THERMOTOGA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RA LA TOUR C., KALTOW H., PORTNER C., CONFALONIERI F., HUBERT R.,
RA DUCQUER M.;
RL SUBMITTED (MAY-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC CC -1- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE
CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.
CC CC -1- CATALYTIC ACTIVITY: ATP-INDEPENDENT BREAKAGE OF SINGLE-STRANDED
CC DNA, FOLLOWED BY PASSAGE AND REJOINING.
CC CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC CC -1- WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA BACKBONE BOND, IT
CC SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN WHICH A TYROSYL OXYGEN
CC IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS AT ONE END OF THE
CC ENZYME-SEVERED DNA STRAND.
CC CC -1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE
CC FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

CC or send an email to license@isb-sib.ch.

CC EMBL: U27841; G881494; -
 DR PROSITE; PS00396; TOPOISOMERASE_I_PROK; 1.
 DR PFAM: PF01131; Topoisom_bac; 1.
 DR HSSP: P06612; 1EC1.
 KW ISOMERASE; TOPOISOMERASE; DNA-BINDING.
 FT ACT_SITE 288 288 DNA CLEAVAGE (BY SIMILARITY).
 SQ SEQUENCE 633 AA; 72694 MW; 9C0D9F68 CRC32;

Query Match 15.1%; Score 61.5; DB 1; Length 633;
 Best Local Similarity 25.4%; Pred. No. 8.5;
 Matches 16; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

QY 18 IKRVSVNIHGRRIDIFASKNFHLQKNTIGTGRRISKLTSKGIASRVYDGIIAAYONPA 77
 DB 96 IARVTNLTGRKNRI-VESEITPRVIRAYKNPREIDMKVRAQLARIIDRLIVGYSLSPV 154
 QY 78 SWK 80
 DB 155 LWR 157

RESULT 10
 RS18_DROME STANDARD; PRT; 152 AA.

AC P41094;
 DT 01-FEB-1995 (REL. 31, CREATED)
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
 DE 40S RIBOSOMAL PROTEIN S18.
 GN RPS18.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOPORPHA; EPHYROIDEA;
 CC DROSOPHILINAE; DROSOPHILA.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-CANTON-S;
 RA MEDLINE: 94215909.
 RX GARWOOD J., LEPEANT J.A.;

RT "The Drosophila melanogaster homolog of ribosomal protein S18.";
 RL GENE 141:231-235(1994).

CC - FUNCTION: INVOLVED IN THE BINDING OF PMET-TRNA AND, HENCE, IN THE
 INITIATION OF TRANSLATION (BY SIMILARITY).
 CC - SIMILARITY: BELONGS TO THE S13P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: L22959; G349256; -
 DR FLYBASE; FBgn0010411; RPS18.
 DR PROSITE; PS00646; RIBOSOMAL_S13; 1.
 DR PFAM: PF00416; S13; 1.
 KW RIBOSOMAL PROTEIN
 SQ SEQUENCE 152 AA; 17611 MW; 72CA0DCB CRC32;

Query Match 15.0%; Score 61; DB 1; Length 152;
 Best Local Similarity 25.6%; Pred. No. 2;
 Matches 21; Conservative 10; Mismatches 29; Indels 22; Gaps 3;

QY 12 PNRVRVIRKSVNIHGR--IDIFASKNFHLQKNTIGTGR-----ISLKLTS 58
 DB 6 PEKFOHLRLIMNTIDGKRKRVGIAMTAIK-----GVGRYSNIIVLAKADVLTAKA 56
 RX 59 GKIASRRVDSIIAAYONPASWK 80

DB 57 GECTEEEDVKVITLISNPLOK 78

RESULT 11
 Y12K_SMSV4 STANDARD; PRT; 110 AA.

AC P36289;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE POSSIBLE 12 KD NUCLEIC ACID-BINDING PROTEIN.
 OS SAN MIGUEL SEA LION VIRUS (SEROTYPE 4) (SMSV 4)
 CC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; CALICIVIRIDAE;
 CC CALICIVIRUS.
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92410750.

RA NEILL J.D.;

RT "Nucleotide sequence of the capsid protein gene of two serotypes of
 RT San Miguel sea lion virus: identification of conserved and non-
 RT conserved amino acid sequences among calicivirus capsid proteins.";
 RL VIRUS RES. 24:211-222(1992).

CC - SIMILARITY: TO FELINE CALICIVIRUS 12 KD PROTEIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M87482; G334889; -
 DR PIR: D48562; D48562.
 SQ SEQUENCE 110 AA; 12566 MW; 24096EF1 CRC32;

Query Match 14.9%; Score 60.5; DB 1; Length 110;
 Best Local Similarity 32.4%; Pred. No. 1.6;
 Matches 11; Conservative 10; Mismatches 12; Indels 1; Gaps 1;

QY 17 FIKRVSVNIHGRRIDIFASKNFHLQKNTIGTGR 50
 DB 10 FLNSVANAYVEGKRLD-ASKGLQLKSRALDTER 42

RESULT 12
 PHS2_RAT STANDARD; PRT; 842 AA.

AC P09812;
 DT 01-MAR-1989 (REL. 10, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE GLYCOGEN PHOSPHORYLASE, MUSCLE FORM (EC 2.4.1.1) (MYOPHOSPHORYLASE).
 GN PYGM.
 OS RATTUS NORVEGICUS (RAT).

CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 CC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
 RN (1)
 RP SEQUENCE OF 2-842 FROM N.A.
 RA HUDSON J.W., HEEFERON K.L., CREERAR M.M.;

RL SUBMITTED (OCT-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN (2)

RP SEQUENCE OF 566-762 FROM N.A.

RX MEDLINE: 86248098

RA OSAMA S., CHIU R.H., MCDONOUGH A., MILLER T.B. JR., JOHNSON G.L.;

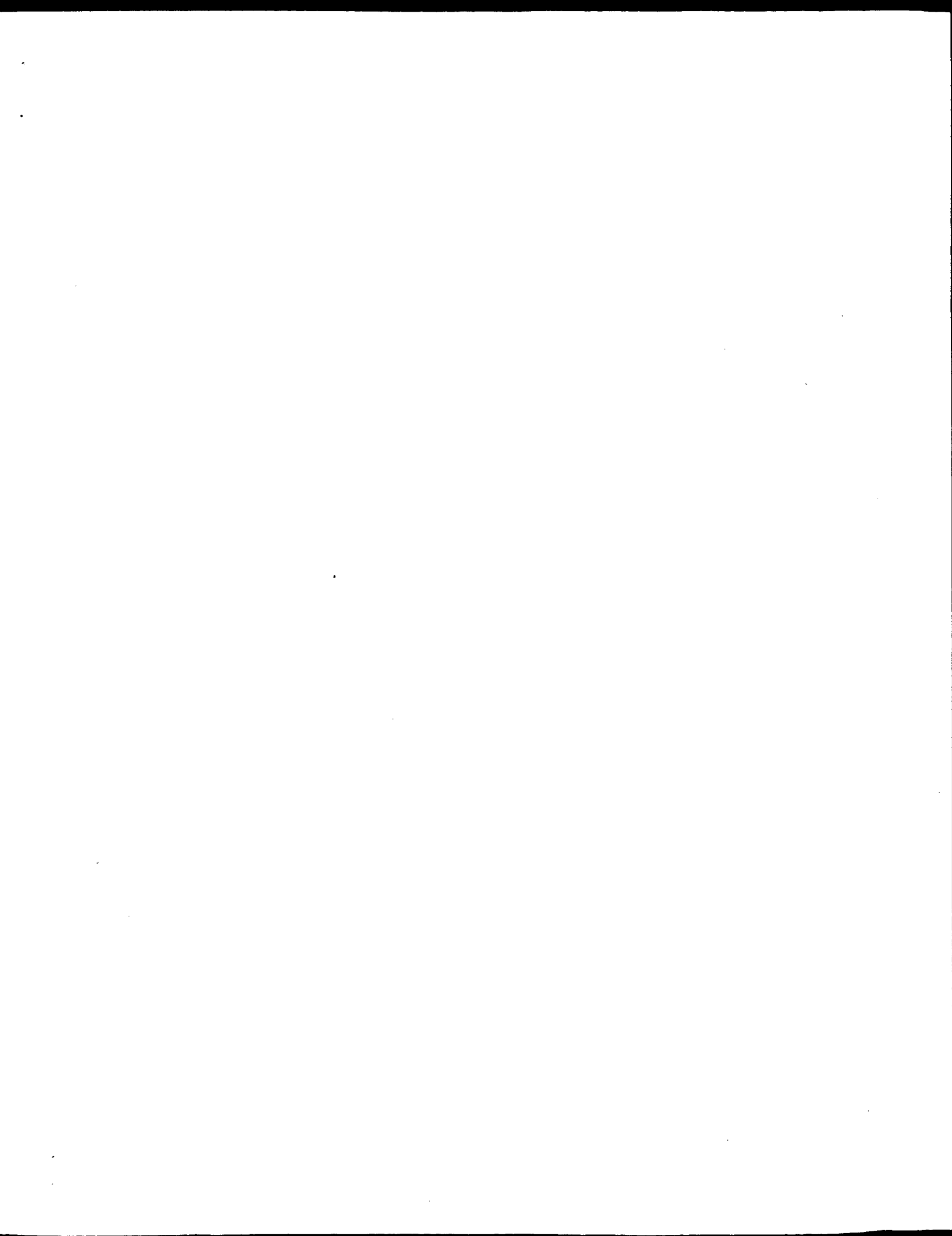
RT "Isolation of partial cDNAs for rat liver and muscle glycogen

RL phosphorylase isozymes.";
 RN FEBS LETT. 202:282-288(1986).

RP SEQUENCE OF 763-842 FROM N.A.
 RX MEDLINE: 86030264.

RA HWANG P.K., SEE Y.P., VINCENTINI A.M., POWERS M.A., FLETTERICK R.J.,
 RA CREAR M.M.:
 RT "Comparative sequence analysis of rat, rabbit, and human muscle
 glyco-gen phosphorylase cDNAs."
 RL EUR. J. BIOCHEM. 152:267-274(1985).
 CC -1- FUNCTION: PHOSPHORYLASE IS AN IMPORTANT ALLOSTERIC ENZYME IN
 CARBOHYDRATE METABOLISM. ENZYMES FROM DIFFERENT SOURCES DIFFER IN
 THEIR REGULATORY MECHANISMS AND IN THEIR NATURAL SUBSTRATES.
 CC HOWEVER, ALL KNOWN PHOSPHORYLASES SHARE CATALYTIC AND STRUCTURAL
 PROPERTIES.
 CC CATALYTIC ACTIVITY: (1,4-ALPHA-D-GLUCOSYL)(N) + ORTHOPHOSPHATE =
 CC (1,4-ALPHA-D-GLUCOSYL)(N-1) + ALPHA-D-GLUCOSE 1-PHOSPHATE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- SUBUNIT: HOMODIMER. DIMERS ASSOCIATE INTO A TETRAMER TO FORM THE
 CC ENZYMATICALLY ACTIVE PHOSPHORYLASE A.
 CC -1- SIMILARITY: BELONGS TO THE GLYCOGEN PHOSPHORYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L10669; G204423; -
 DR EMBL: X03032; G818028; -
 DR PIR: B24200; G24200.
 DR PIR: C23093; C23093.
 DR PROSITE: PS00102; PHOSPHORYLASE; 1.
 DR PFAM: PF00343; phosphorylase; 1.
 DR HSP: P00489; 1PYG.
 DR TRANSFERASE: GLYCOSYLTRANSFERASE; CARBOHYDRATE METABOLISM;
 DR GLYCOGEN METABOLISM; ALLOSTERIC ENZYME; PYRIDOXAL PHOSPHATE;
 DR PHOSPHORYLATION.
 KW MOD_RES 15 15
 FT BINDING 76 76
 FT SITE 109 109
 FT SITE 143 143
 FT ACT_SITE 156 156
 FT BINDING 681 681
 FT CONFLICT 640 641
 FT CONFLICT 724 724
 FT CONFLICT 766 766
 SQ SEQUENCE 842 AA; 97273 MW; 32741482 CRC32;
 Query Match 14.7%; Score 59.5; DB 1; Length 842;
 Best Local Similarity 29.7%; Pred. No. 20;
 Matches 22; Conservative 9; Mismatches 30; Indels 13; Gaps 3;
 QY 10 FGNNRRVFIRKRVSVNIIHGRIDIFA-----SKNHLQKNTIGTGRRLSLKI-TSG 59
 Db 751 FSPKQPDLEFDIVNMVHHDFKVFADYEYIKQDKVSELYKNPRETRMVIKRNATSG 810
 QY 60 KIASRRVDGIITAY 73
 Db 811 KFS-----DRIIAY 821
 RESULT 13
 FKX2 YEAST
 ID FKX2 YEAST STANDARD: PRI: 862 AA.
 AC P41813;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE FORK HEAD PROTEIN HOMOLOG 2.

GN FKX2 OR YNL068C OR N2403 OR YNL2403C.
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C.
 RC STRAIN-S288C.
 RA ZHU G., DAVIS T.N.;
 RL SUBMITTED (XXX-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / FY1679;
 RX MEDLINE: 96267764.
 RA POEHLMANN R., PHILIPSEN P.;
 RT "Sequencing a cosmid clone of *Saccharomyces cerevisiae* chromosome XIV
 RT reveals 12 new open reading frames (ORFs) and an ancient duplication
 RT of six ORFs."
 RL YEAST 12:391-402(1996).
 RN [3]
 RP SEQUENCE OF 1-440 FROM N.A.
 RC STRAIN-S288C / FY1676;
 RX MEDLINE: 96021608.
 RA BERGERZ P., DOIGNON F., CROUZET M.;
 RT "The sequence of a 44 420 bp fragment located on the left arm of
 RT chromosome XIV from *Saccharomyces cerevisiae*."
 RL YEAST 11:967-974(1995).
 RN [4]
 RP ERRTATUM.
 RA MEDLINE: 97060022.
 RA BERGERZ P., DOIGNON F., CROUZET M.;
 RL YEAST 12:297-297(1996).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 FHA DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L38850; G623608; -
 DR EMBL: X86470; G791119; -
 DR EMBL: Z71343; E328836; -
 DR EMBL: Z71344; E239903; -
 DR EMBL: U12141; G1098481; -
 DR SGD: L0002608; FKX2.
 DR PROSITE: PS00657; FORK_HEAD_1; 1.
 DR PROSITE: PS00658; FORK_HEAD_2; 1.
 DR PROSITE: PS50006; FHA_DOMAIN; 1.
 DR PROSITE: PS50039; FORK_HEAD_3; 1.
 DR PFAM: PF00250; Fork_head; 1.
 DR PFAM: PF00498; FHA; 1.
 DR DNA-BINDING; NUCLEAR PROTEIN.
 KM DOMAIN 83 152
 FT DNA_BIND 339 430
 FT DNA_BIND 339 430
 SQ SEQUENCE 862 AA; 94374 MW; 83DA8394 CRC32;
 Query Match 14.5%; Score 59; DB 1; Length 862;
 Best Local Similarity 30.6%; Pred. No. 24;
 Matches 22; Conservative 11; Mismatches 31; Indels 8; Gaps 4;
 QY 11 GPNRRVFIRKRVSVNIIHGRIDIFASKNFHLOKNTIG--TGRISLKLSTGKIASRRVDG 68
 Db 71 GPNRTYYVKLEYSI--GRNTD---PLNSALQENSDGVKNSYKRVNIDLGPAKVSRRK-HA 124
 QY 69 ITAAYQNPASWK 80
 Db 125 IIKYNNMGWE 136



GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 1999, 13:02:00 ; Search time 45.11 Seconds
(without alignments)
109.144 Million cell updates/sec

Title: US-09-142-524-1

Perfect score: 406
Sequence: 1 MKVTVAFNQFGPNRRVFIKR.....IASRRVDGITAAYONPASMK 80

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOIST:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	18.2	375	10 Q96385	Q96385 chameacypar
2	73	18.0	367	10 Q96385	Q96385 chameacypar
3	72	17.7	334	5 Q46197	Q46197 dirosophila
4	68	16.7	330	10 Q65456	Q65456 arabidopsis
5	66	16.3	350	2 P73865	P73865 synchocyst
6	66	16.3	394	10 Q65457	Q65457 arabidopsis
7	65	16.0	390	10 Q65388	Q65388 arabidopsis
8	65	16.0	498	12 Q88525	Q88525 turkey hep
9	64	15.8	222	8 Q21280	Q21280 reclinomona
10	63.5	15.6	533	5 Q01617	Q01617 caenorhabdi
11	61.5	15.1	102	2 Q66004	Q66004 mycoplasma
12	61	15.0	173	2 Q47446	Q47446 escherichia
13	60	14.8	726	4 Q15152	Q15152 homo sapien
14	60	14.8	725	4 Q13835	Q13835 homo sapien
15	60	14.8	275	5 Q44584	Q44584 caenorhabdi
16	60	14.8	587	5 Q43956	Q43956 tetrahymena
17	60	14.8	154	5 Q18240	Q18240 caenorhabdi
18	60	14.8	295	6 Q28875	Q28875 bos taurus
19	60	14.8	727	6 Q28161	Q28161 bos taurus
20	60	14.8	627	10 Q64766	Q64766 arabidopsis
21	60	14.8	257	10 Q96387	Q96387 nicotiana t
22	59.5	14.7	254	1 Q30306	Q30306 archaeglob
23	59.5	14.7	269	10 Q48980	Q48980 hordeum vul
24	59	14.5	152	3 Q94754	Q94754 schizosacch
25	58.5	14.4	380	11 P97877	P97877 rattus norv
26	58	14.3	729	1 P95924	P95924 sulfolobus
27	58	14.3	1396	2 Q05450	Q05450 mycobacteri
28	58	14.3	578	2 Q92K18	Q92K18 helicobacte
29	58	14.3	661	5 Q21394	Q21394 caenorhabdi

30	58	14.3	728	11 P97350	P97350 mus musculu
31	57.5	14.2	103	2 Q51514	Q51514 borrelia bu
32	57.5	14.2	314	2 Q95G17	Q95G17 borrelia
33	57.5	14.2	796	5 Q17532	Q17532 caenorhabdi
34	57.5	14.2	834	5 Q19783	Q19783 caenorhabdi
35	57	14.0	616	1 Q51714	Q51714 pyrococcus
36	57	14.0	314	2 Q92ANS	Q92ANS comamonas s
37	57	14.0	593	3 Q60102	Q60102 schizosacch
38	57	14.0	656	9 Q38355	Q38355 lactococcus
39	56.5	13.9	254	2 P74730	P74730 synchocyst
40	56.5	13.9	339	2 Q87960	Q87960 borrelia bu
41	56.5	13.9	538	5 Q16342	Q16342 caenorhabdi
42	56.5	13.9	563	5 Q45758	Q45758 caenorhabdi
43	56.5	13.9	431	10 Q23017	Q23017 arabidopsis
44	56.5	13.9	108	12 Q09783	Q09783 san miguel
45	56	13.8	492	4 Q14926	Q14926 homo sapien

ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT;	375 AA.
ID Q96385			
AC Q96385;			
DT 01-FEB-1997 (TREMBLrel. 02, Created)			
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)			
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)			
DE CHA01 PRECURSOR.			
OS Chameacyparis obtusa.			
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;			
OC Cupressaceae; Chameacyparis.			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE-POLLEN.			
RX MEDLINE: 96265194.			
RA SUZUKI M., KOMIYAMA N., ITOH M., ITOH H., SONE T., KUNO K., TAKAGI I.,			
RA OHTA N.;			
RT "Purification, characterization and molecular cloning of Cha o 1," a			
RT major allergen of Chameacyparis obtusa (Japanese cypress) pollen.";			
RL Mol Immunol. 33:451-460(1996).			
DR EMBL: D45404; BAA08246.1; -			
DR MENDEL: 7626; CHAOB:1088;1.			
DR PFAM: PF00544; pec_lyase; 1.			
DR SIGNAL.			
FT SIGNAL.			
KW CHAIN			
FT CHAIN			
SO SEQUENCE			
375 AA; 40258 MW; A0981492 CRC32;			
Query Match	18.2%;	Score 74;	DB 10;
Best Local Similarity	59.3%;	Pred. No. 0.51;	Length 375;
Matches	16;	Conservative	2;
Mismatches	9;	Indels	0;
Gaps	0;		
QY 1 MKVTVAFNQFGPNRRVFIKRVSNVIT 27			
DB 234 MKVTVAFNQFGPNRRVFIKRVSNVIT 260			
RESULT 2			
Q92NUT			
ID Q92NUT;			
AC Q92NUT;			
DT 01-MAY-1999 (TREMBLrel. 10, Created)			
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)			
DE POLLEN MAJOR ALLERGEN 1-2.			
OS Juniperus ashei (Ozark white cedar).			
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;			
OC Cupressaceae; Juniperus.			
RN [1]			

RP SEQUENCE FROM N.A.
 RA MIDORO-HORIUTI T.M., GOLDBLUM R.M., KUROSKY A., WOOD T.G.,
 RA BROOKS E.G.;
 RT "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
 RT allergen, Jun a 1.1";
 DR EMBL: AF106663; AAD03609.1; -
 DR EMBL: AF106662; AAD03608.1; -
 SQ SEQUENCE 367 AA; 39824 MW; 4C2DB630 CRC32;

Query Match 18.0%; Score 73; DB 10; Length 367;
 Best Local Similarity 55.6%; Pred. No. 0.65;

Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTVAENQFGPNRRVFIRKRVSVIHH 27
 DB 234 MKVTVAENQFGPNRRVFIRKRVSVIHH 260

RESULT 3
 ID 046197 PRELIMINARY; PRT; 234 AA.
 AC 046197;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, last annotation update)
 DE ACCESSORY GLAND PROTEIN ACP29AB.
 GN ACP29AB.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
 OC Drosophilidae; Drosophila.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CANTON S; TISSUE-ACCESSORY GLAND;
 RA WOLFNER M.F., HARADA H.A., BERTRAM M.J., STELICK T.J., KRAUS K.W.,
 RA KALB J.M., LUNG Y.O., NEUBAU D.M., PARK M., TRAM U.;
 RL Insect Biochem. Mol. Biol. 27:0-0(1997).
 DR EMBL: U85758; AAB96382.1; -
 DR FLXBASE: FB91001583; ACP29AB.
 DR PFAM: PF00059; Lectin_c1.
 SQ SEQUENCE 234 AA; 27173 MW; F674589C CRC32;

Query Match 17.7%; Score 72; DB 5; Length 234;
 Best Local Similarity 25.9%; Pred. No. 0.49;

Matches 21; Conservative 19; Mismatches 31; Indels 10; Gaps 3;

QY 9 OFGPNRRVFIRKRVSVIHHGRIDIFASKNFHLQKNTIGT--GRRISLKLTSKRIA 62
 DB 88 QLOPLKIRHSHASNNIKASNNIKMRFEVSRHRIENKLMQTFEAYVTCRKNGHILA 147
 OY 63 S---RRVDGIIAAYONPASW 79
 DB 148 NIDDEKELDGIALLAPNNISYW 168

RESULT 4
 ID 065456 PRELIMINARY; PRT; 394 AA.
 AC 065456;
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, last annotation update)
 DE PEPTIDE LIASE LIKE PROTEIN.
 GN FLN20.180.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN [1]

RP SEQUENCE FROM N.A.
 RA BEVAN M., WEDLER H., WAMBUIT R., BANCROFT I., MEWES H.W., MAYER K.,
 RA SCHUELLER C.;
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU ARABIDOPSIS SEQUENCING PROJECT;
 RA Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL022140; CAA18111.1; -
 DR PFAM: PF00544; pec_lyase; 1.
 KM lyase.
 SQ SEQUENCE 394 AA; 43476 MW; 95399178 CRC32;

Query Match 16.7%; Score 68; DB 10; Length 394;
 Best Local Similarity 28.6%; Pred. No. 2.8;

Matches 24; Conservative 7; Mismatches 27; Indels 26; Gaps 3;

QY 1 MKVTVAENQFGPNRRVFIRKRVSVIHH--GRID-----IFASKNFHLQKN 44
 DB 256 MKVTVAENQFGPNRRVFIRKRVSVIHH--GRID-----IFASKNFHLQKN 44

QY 45 TIGTGRIISLKLTSKRIASRRVDG 68
 DB 316 -----KSNKRVTKREVKG 329

RESULT 5
 ID P73865 PRELIMINARY; PRT; 350 AA.
 AC P73865;
 DT 01-FEB-1997 (TREMblrel. 02, Created)
 DT 01-FEB-1997 (TREMblrel. 02, last sequence update)
 DT 01-JAN-1999 (TREMblrel. 09, last annotation update)
 DE SENSOR TRANSDUCTION HISTIDINE KINASE.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PCC6803;
 RA TABATA S.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PCC6803;
 RX MEDLINE: 97061201.
 RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
 RA MIYATA N., HIROGAWA M., SUGIURA M., SASAKOTO S., KIMURA T.,
 RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
 RA SHIMO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M.,
 RA TABATA S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL: D90910; BAA17927.1; -
 DR PFAM: PF00512; signal; 1.
 SQ SEQUENCE 350 AA; 39445 MW; EEPF2A0B CRC32;

Query Match 16.3%; Score 66; DB 2; Length 350;
 Best Local Similarity 39.5%; Pred. No. 4.2;

Matches 17; Conservative 7; Mismatches 15; Indels 4; Gaps 2;

QY 14 RRVFIKRVSVIHH--GRIDI-FASKNFHLQKNTIGRRT 52
 DB 246 GRVFNILITAINHSPRGKRVETISLTSKNHFOYIVDEGRGI 288

RESULT 6
 ID 065457 PRELIMINARY; PRT; 394 AA.
 AC 065457;

[illegible]

RESULT	7	
0653388		
ID	0653388	PRELIMINARY; PRT; 390 AA.
AC	0653388;	
DT	01-AUG-1998 (TREMBLrel. 07, Created)	
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)	
DT	01-MAY-1999 (TREMBLrel. 10, Last annotation update)	
DE	FlzF1.22 PROTEIN.	
GN	FlzF1.22.	
OS	Arabidopsis thaliana (Mouse-ear cross).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;	
OC	core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae;	
OC	Arabidopsis.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-CV. COLUMBIA:	
RA	YVSCOSKIA V.S., OSBORNE B.I., SCHWARTZ J.R., TORIUMI M., YU G.,	
RA	KWAN A., OJI O., LUD S., BOEHLER E., CONWAY A.B., CONWAY A.R.,	
RA	DEMAR K., FENG J., KIM C., KURTZ D., PALM C.J., LI Y., SHINN P.,	
RA	SUN H., DAVIS R.W., ECKER J.R., FEDERSPIEL N.A., THEOLOGIS A.,	
RL	Submitted (MAY-1998) to the EMBL/GenBank/DBD databases.	
DR	EMBL; ACC002143; AAC17625.1; -	
DR	FEAM; PF005444; pec_lyase; 1.	
SO	SEQUENCE 390 AA; 43354 MW; E0F322FB4 CRC32;	

	Query Match	16.0%	Score 65;	DB 10;	Length 390;
	Best Local Similarity	25.0%	Pred. No. 6, 2;		
Matches	25;	Conservative	1;	Mismatches	29;
				Indels	28;
				Gaps	5;
Qy	1 MKVTAAFNPFNOGGRNRVRFLKRSNYIIH-----GRRD--IFASKNFHLOKN	44			
Db	246 MGVTLAIFAHFGGLLEKRPFRBRYAVANNRREKKOMTAIGSGADPTFSSGNF----	302			

RESULT	8			
088525				
ID	088525	PRELIMINARY;	PTI;	498 AA.
AC	088525;			
DT	01-NOV-1996 (T-EMBLrel. 01, Created)			
DT	01-NOV-1996 (T-EMBLrel. 01, Last sequence update)			
DT	01-NOV-1998 (T-EMBLrel. 08, Last annotation update)			
DE	GLYCOPROTEIN HOMOLOGUE E.			
OS	Turkey Herpesvirus.			
NC	Viruses; dsDNA_viruses, no RNA stage; Herpesviridae.			
OC	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE: 94014999.			
RA	ZELINK V., DAREIL R., AUDONNET J., SMITH G.D., RIVIERE M.,			
RA	PASTOREK J., ROSS N.L.J.;			
RT	"The complete sequence and gene organization of the short unique			
RT	region of herpesvirus of turkeys."			
RL	J. Gen. Virol. 74:2151-2162(1993).			
DR	EMBL: x68653. CAA48619.1;			
SO	SEQUENCE 498 AA; 55059 MW; 0E20E22F CRC32;			

```

Query Match 16.0% Score 65; DB 12; Length 498;
Best Local Similarity 31.6%; Pred. No. 8, 4
Matches 18; Conservative 10; Mismatches 25; Indels 4; Gaps 2

18 IKRVSVIIIGRRIDIPASKNHFLOKNT-ITGGRISLKTSGKIASRVDYITAAV 73
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
92 ILRVDFVYNSSLDIRIASSQFHPIPNVITDGTGKELLTFN--ATVSGIAYVTR 145

```

```

RESULT      9
ID 021280
AC 021280; PRELIMINARY; PRT; 222 AA.
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE APP-BINDING SUBUNIT OF ABC TRANSPORTER FOR CYTOCHROME C1.
GN YEW.
OS Reclinomonas americana.
OC Mitochondrion.
OC Eukaryota; Reclinomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC50394.
RX MEDLINE: 97311393.
RA LANG B.F., BURGER G., O'KELLY C.J., CEDERGREEN R., GOLDING G.B.,
RA LEMIEUX C., SANKOFF D., TREMBL M., GRAY M.W.;
RT "An ancestral mitochondrial DNA resembling a eubacterial genome in
RL Nature 387:493-497(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC50394.
RA LANG B.F., BURGER G.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF007261; AD11907.1; -.
DR PFM: PF00005; ABC_tran; 1.
KW ATP-binding; Mitochondrion.
SQ SEQUENCE 222 AA; 25502 MW; D81B02D7 CRC32;

```

Query Match 15.8%; Score 64; DB 8; Length 222;
Best Local Similarity 30.8%; Pred. No. 4.1;
Matches 20; Conservative 14; Mismatches 19; Indels 12; Gaps 3;
QY 18 IKRSNVIIH---GRID--IFASKNFHLQNTY-----GICGRRLSLTSGKIASRR 65

DB 3 LTRQWLTHTNITGIRSKNITIFONINFSLEKGSLEPIQSGNSGRTTLTKITISGLPFSQ 62

QY 66 VDGII 70

DB 63 GDIII 67

RESULT 10

ID 001617 PRELIMINARY; PRT; 533 AA.

AC 001617; 01-JUN-1998 (TREMblrel. 06, Created)

DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)

DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)

DE HYPOTHETICAL 60.7 KD PROTEIN T19H12.1 IN CHROMOSOME V PRECURSOR.

GN T19H12.1.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae; Rhabditina; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA DAVIDSON S.;

RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

DR EMBL; 097009; AAC69026.1; -

DR WORMPEP; T19H12.1; CE13752.

DR PFAM; PF00201; UDPGT; 1.

AM Hypothetical protein; Signal; Transmembrane.

FI SIGNAL 1 18 POTENTIAL.

FT CHAIN 19 533 HYPOTHETICAL PROTEIN T19H12.1.

FT TRANSMEM 152 172 POTENTIAL.

FT TRANSMEM 260 280 POTENTIAL.

DR TRANSMEM 495 515 POTENTIAL.

SO SEQUENCE 533 AA; 60726 MW; 008AAFTD CRC32;

Query Match 15.6%; Score 63.5; DB 5; Length 533;

Best Local Similarity 28.2%; Pred. No. 14;

Matches 19; Conservative 14; Mismatches 17; Indels 15; Gaps 4;

QY 10 FGPNRRVEIKRVSNI-HGRRIDIFASKNFHLQNTIGTGRISLTKTSKIASRVVG 68

DB 28 FGYSKAFVVSOLANTIDHGHNVTF-QPFH-----IALKNTDGLINKNIE- 74

QY 69 IIAAY 73

DB 74 IINYY 78

RESULT 11

ID 066004 PRELIMINARY; PRT; 102 AA.

AC 066004; 01-AUG-1998 (TREMblrel. 07, Created)

DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)

DT 01-AUG-1998 (TREMblrel. 07, Last annotation update)

DE MANNITOL-1-PHOSPHATE DEHYDROGENASE.

GN MTLID.

OS Mycoplasma bovine group 7.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites; Mycoplasmataceae; Mycoplasma.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-PG50;

RA MEDLINE; 98439347.

RA FREY J., CHENG X., MONNERAT M.-P., ABDO E.-M., KRAMINKLER M., BOLSENE G., NICOLET J.;

RT "Genetic and serological analysis of the immunogenic 67-kDa lipoprotein of Mycoplasma sp. bovine group 7.";

RL Res. Microbiol. 149:55-64(1998).

DR EMBL; 067071; AAC06132.1; -

SO SEQUENCE 102 AA; 12119 MW; 28FIAC19 CRC32;

Query Match 15.1%; Score 61.5; DB 2; Length 102;

Best Local Similarity 24.3%; Pred. No. 3.2;

Matches 18; Conservative 15; Mismatches 36; Indels 5; Gaps 1;

QY 2 KVTAFNPGPNRRVEIKRVSNI-HGRRIDIFASKNFHLQNTIGTGRISLTKTSKI 61

DB 25 KYNLSNLENDYKIQILRLNSPFIKDDLRKLVNTELEKLSKN-----EELITILDYAVY 79

QY 62 ASRRVDGIIAAYQ 75

DB 80 SNLKHDTLLSLYQN 93

RESULT 12

ID 047446 PRELIMINARY; PRT; 173 AA.

AC 047446; 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)

DE PAPE PROTEIN.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE; 88169520.

RA LUND B., LINDBERG F., NORMARK S.;

RT "Structure and antigenic properties of the tip-located P pilus proteins of uropathogenic Escherichia coli.";

RL J. Bacteriol. 170:1887-1894(1988).

DR EMBL; M20181; AAA24291.1; -

DR PFAM; PF00419; Fimbria; 1.

SO SEQUENCE 173 AA; 18450 MW; E09D027D CRC32;

Query Match 15.0%; Score 61; DB 2; Length 173;

Best Local Similarity 27.0%; Pred. No. 7;

Matches 20; Conservative 11; Mismatches 35; Indels 8; Gaps 3;

QY 1 MKVTY-AFNGPGRNRRVEIKRVSNI-HGRRIDIFASKNFHLQNTIGTGRISLTKTSK 59

DB 79 MKVTYATNTY--NNALIVNTSNTSSDGLVLYLNS-----MAGNIGTATLTGPTPTPG 131

QY 60 KIASRRVDGIIAAY 73

DB 132 KITGNMADRTISLH 145

RESULT 13

ID 015152 PRELIMINARY; PRT; 726 AA.

AC 015152; 000645; 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)

DE PLAKOPHILIN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RC SCHMIDT A., HANS H.W., SCHAEFER S., NOBER U.A., ZIMBELMANN R., FRANK W.W.;

RA Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.

RL [2]

RP SEQUENCE FROM N.A.

RA ZIMBELMANN R.;

RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.
 RA BOSCH A.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN
 RN (4)
 SEQUENCE FROM N.A.
 RA SCHMIDT A., LANGBEIN L., RODE M., PRAETZEL S., FRANK W.W.;
 RL J. Mol. Biol. 0:0-0(0).
 DR EMBL; 234974; CAAB4426.1; -
 DR EMBL; 273678; CAA8022.1; -
 DR PFAM; PF00514; Armadillo_seg; 3.
 SO SEQUENCE 726 AA; 80496 MW; A92E30E9 CRC32;

Query Match 14.8%; Score 60; DB 4; Length 726;
 Best Local Similarity 30.8%; Pred. No. 52;
 Matches 12; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTGRISIKLTSGRKIASRRVDGI 69
 DB 293 VDLRSPNQNVQAAAGALRNLFVSTNKLSTRONGI 331

RESULT 14
 ID 013835 PRELIMINARY; PRT; 725 AA.
 AC 013835;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE BAND-6-PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EPIDERMIS;
 RX MEDLINE; 95074299.
 RA HATZFELD M., KRISTJANSSON G.I., PLESSMANN U., WEBER K.;
 RT "Band 6 protein, a major constituent of desmosomes from stratified
 epithelia, is a novel member of the armadillo multigene family."
 RL J. Cell Sci. 107:2259-2270(1994).
 DR EMBL; X79293; CAA55881.1; -
 DR PFAM; PF00514; Armadillo_seg; 3.
 SO SEQUENCE 725 AA; 80433 MW; 270BE242 CRC32;

Query Match 14.8%; Score 60; DB 4; Length 725;
 Best Local Similarity 30.8%; Pred. No. 52;
 Matches 12; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTGRISIKLTSGRKIASRRVDGI 69
 DB 292 VDLRSPNQNVQAAAGALRNLFVSTNKLSTRONGI 330

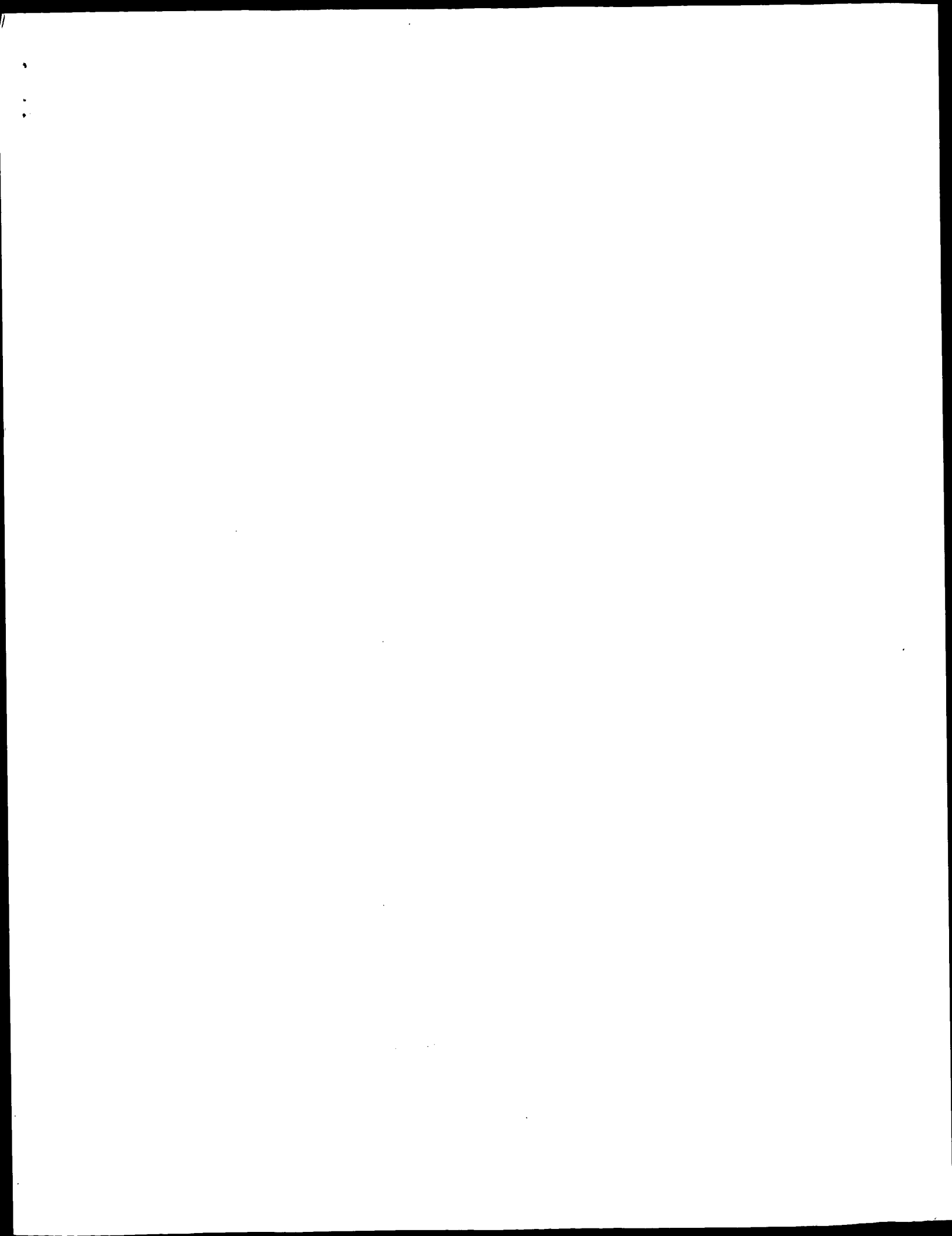
RESULT 15
 ID 044584 PRELIMINARY; PRT; 275 AA.
 AC 044584;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE F4867.4 PROTEIN.
 GN F4867.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Secernentea; Rhabdita; Rhabditida;
 CC Rhabditina; Rhabditidae; Peloderinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE; 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,

RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans."
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA CLARKE K., WOHLDMANN P., HARRISON M.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF039044; AF047944.1; -
 SO SEQUENCE 275 AA; 32128 MW; E9B82465 CRC32;

Query Match 14.8%; Score 60; DB 5; Length 775;
 Best Local Similarity 34.3%; Pred. No. 16;
 Matches 12; Conservative 10; Mismatches 9; Indels 4; Gaps 1;

QY 36 SKNFHLQKNTIGTGRISIKLTSGRKIASRRVDGI 70
 DB 230 SKRFYOKLQISTGERVLIV---KLSADTIEGVV 260

Search completed: September 29, 1999, 13:36:45
 Job time: 2085 sec



OM of: US-09-142-524-1 to: GenEmbl.* out-format : pfs

Date: Sep 29, 1999 1:59 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-1998 Compugen Ltd.

Command line parameters:

-MODEL-frame1_p2n.model -DEV-xlp
-O/cgn2/_/us09142524/runat_29091999.125958.12782/arp_query.fasta.1
-UB-GenEmbl -OFMT=fastap -SUFFIX=rge -GAPD=12.000 -GAPEXT=4.500
-MATCH=0.100 -LOPCL=0.000 -LOPEXT=0.000 -GAPD=6.000
-MATCH=0.050 -XGAPD=10.000 -XGAPEXT=0.500 -GAPD=6.000
-MATCH=7.000 -YGAPD=10.000 -YGAPEXT=0.500 -DELOP=6.000
-MATCH=7.000 -STAR=1 -MATRIX-biosum62 -TRANS-human40.cdi
-157-15 -DOCALLIGN=200 -THR_SCORE=esscore -ALIGN=15 -MODE=LOCAL
-OFMT=pfs -NORW=stat -USER=US09142524 -NCPU=6 -ICPU=3 -WAIT
-THREADS=1

Header information block:

Query: US-09-142-524-1
Query length: 80
Database: GenEmbl.*
Database sequences: 679419
Database length: 1590154680
Search time (sec): 583.430000

Score_list:

Sequence	Strd Orig	ZScore	Escore	len	Documentation
em_pat:E09607	+	99.00	202.96	0.0023	1545
em_pat:E09664	+	99.00	202.96	0.0023	1545
gb_p11:CPOCRV11	+	99.00	202.96	0.0026	1708
em_pat:E10716	+	99.00	201.90	0.0026	1733
gb_p11:CPOCRV11	+	99.00	201.91	0.0027	1790
em_pat:E11677	+	74.00	149.69	2.13	1062
gb_p11:E11678	+	74.00	149.16	2.28	1125
em_pat:E11678	+	74.00	148.13	2.61	1260
gb_p11:D45404	+	74.00	147.85	2.93	792
em_pat:E11679	+	72.00	147.24	2.93	516
gb_p11:D45404	+	70.00	147.08	2.99	525
gb_p12:AF106662	+	73.00	147.07	2.99	1104
gb_p12:AF106663	+	73.00	147.07	2.99	1104
gb_om:AB000911	+	70.00	147.03	3.01	528
gb_p12:HSPRS18	+	70.00	146.67	3.15	549
em_pat:E11681	+	74.00	145.00	3.31	1542
em_pat:E11681	+	74.00	145.00	3.31	1542
gb_p11:CPOCRV11	+	71.00	141.35	6.22	1256
gb_p11:CPOCRV11	+	71.00	141.35	6.23	1257
gb_p11:CPOCRV11	+	71.00	141.29	6.27	1264
gb_p11:CPOCRV11	+	71.00	140.92	6.58	1317
gb_p12:HSPRS18	+	60.50	137.33	10.43	145
gb_p12:HSPRS18	+	62.50	134.44	15.10	326
gb_om:AF051365	+	63.50	130.67	24.50	630
gb_p12:HSPRS18	+	61.00	127.63	36.19	473
gb_om:ST037027	+	59.50	126.46	42.21	372
gb_p12:HSPRS18	+	66.00	126.30	42.88	1883
gb_p12:HSPRS18	+	61.00	125.96	44.79	567
gb_om:SSU15448	+	59.00	123.92	58.20	432
gb_p12:HSPRS18	+	61.00	123.56	60.96	732
gb_p12:HSPRS18	+	61.00	123.56	60.96	732
gb_om:AF034264	+	63.50	122.96	65.80	1400
gb_om:AF034264	+	59.50	122.96	68.14	559
gb_om:AF034264	+	64.50	122.28	71.84	2015
gb_p12:HSPRS18	+	57.50	121.82	76.24	372
gb_p12:HSPRS18	+	57.50	121.41	80.32	395
gb_p12:HSPRS18	+	65.00	121.05	84.12	2608
gb_p12:HSPRS18	+	65.00	121.05	84.12	2608
gb_p11:YSCSECI7P	+	62.50	121.04	84.19	1406

gb_pat:A67730 + 50.00 121.01 84.53 64 ! A67730 Sequence 60 from Paten
gb_p11:ECOPAPER + 58.00 120.74 87.47 477 ! G48233 SHGC-60275 Human Homo
gb_p11:ECOPAPER + 62.00 120.74 87.52 1284 ! M13239 E.Coli papE and papF

seq_name: em_pat:E09607

seq_documentation_block:

ID E09607 standard: RNA: UNC; 1545 BP.

AC E09607:

SV E09607.1

NI d1107944

DT 08-OCT-1997 (Rel. 52, Created)

DT 08-OCT-1997 (Rel. 52, Last updated, Version 1)

DE CDNA encoding polypeptide causing pollen allergy.

XX JP 1995170986-A/1.

XX unclassified

XX unclassified

XX [1]

XX 1-1545

RA Namba M., Torigoe K., Kurimoto M.;

RT "DNA CODING POLYPEPTIDE, RECOMBINANT DNA CONTAINING THE DNA AND

RL TRANSFORMANT";

RL Patent number JP 1995170986-A/1, 11-JUL-1995.

RL HAYASHIBARA BIOCHEM LAB INC.

OS Cryptomeria (cedar)

PN JP 1995170986-A/1

CC PD 11-JUL-1995

CC PF 20-DEC-1993 JP 1993344596

CC PR 05-NOV-1993 JP 93P 299151

CC PI NAKAMA MOTOJI, TORIGOE KAKUJI, KURIMOTO MASASHI

CC PC C12N15/09, C12N1/21, G01N33/50, A61K39/36, G01N33/53, G01N33/53,

CC PC C12N1/21;

CC CC strandedness: Double;

CC CC topology: linear;

CC CC hypothetical: No;

CC CC anti-sense: No;

CC FH key Location/Qualifiers

CC FH 1.1545

CC FT source

CC FT 1.1545

CC FT /db_xref="taxon:32644"

CC FT /organism="unidentified"

CC FT key Location/Qualifiers

CC FT 1.1545

CC FT source

CC FT 1.1545

CC FT /db_xref="taxon:32644"

align_scores: 99.00

Ratio: 5.211

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-09-142-524-1 x E09607

Align seg 1/1 to: E09607 from: 1 to: 1545

31 ILEASPILEPHEALASERLYSASNPHEHISLENGHLYSASNPHTLLEGI 47
 |||||||
 706 ATTGATATCTTTCATCTATAAACTTTCATCTTCATAAGACACGATAG 755

47 ythrngly 49
 |||||||
 756 AACAGGG 762

seq_name: em_pat:E09664

seq_documentation_block:
 ID E09664 standard; RNA; UNC; 1545 BP.

AC E09664;
 XX
 SV E09664.1
 XX
 NT d1108001

07-OCT-1997 (Rel. 52, Created)
 07-OCT-1997 (Rel. 52, Last updated, Version 1)

CUWA encoding a novel protein which induces hay fever.

JP 1995188289-A/1.

unclassified.
 OC
 OS
 XX
 KW
 AA
 DT
 YY

[1]
 RN 1-1545

RA Torigoe K., Namba M., Kurimoto M.;

RT "POLYPEPTIDE, ITS PRODUCTION AND USE";

RL Patent number JP 1995188289-A/1, 25-JUL-1995.

RL HAYASHIBARA BIOCHEM LAB INC.

XX
 CC OS Cryptomeria sp.

CC PN JP 1995188289-A/1

CC PD 25-JUL-1995

CC PF 27-DEC-1993 JP 1993346814

CC PI TORIGOE KAKUI, NAMBA MOTOJI, KURIMOTO MASASHI

CC PC C07K14/415,A61K39/36,C12N15/09,C12P21/02,(C12P21/02,C12R1.19);

CC CC strandedness: Double;

CC CC topology: Linear;

CC FH Key Location/Qualifiers

CC FT source 1. 1545

CC FT CDS /organism="Cryptomeria sp."
 /tissue-type="pollen"
 1. 1545
 /product="hay fever-inducing protein"

CC FH Key Location/Qualifiers

CC FT source 1. 1545

CC FT /db_xref="taxon:32644"

CC FT /organism="unidentified"

XX Sequence 1545 BP; 508 A; 281 C; 354 G; 402 T; 0 other;

alignment_scores:

Quality: 99.00 Length: 19

Ratio: 5.211 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-142-524-1 x E09664 ..

Align seg 1/1 to: E09664 from: 1 to: 1545

31 ILEASPILEPHEALASERLYSASNPHEHISLENGHLYSASNPHTLLEGI 47
 |||||||

706 ATTGATATCTTTCATCTATAAACTTTCATCTTCATAAGACACGATAG 755

47 ythrngly 49
 |||||||
 756 AACAGGG 762

seq_name: gb_p11:CP0CRXJ11

seq_documentation_block:

LOCUS CP0CRXJ11 1708 bp mRNA PLN 07-FEB-1999

DEFINITION Japanese cedar mRNA for Cry j II allergen, complete cds.

ACCESSION D29772

NID 9506857

VERSION D29772.1 GI:506857

KEYWORDS Cry j II; allergen.

SOURCE Cryptomeria japonica cDNA to mRNA.

ORGANISM Cryptomeria japonica

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

euphylliphytes; Spermatophyta; Coniferopsida; Coniferales;

Taxodiaceae; Cryptomeria.

1 (bases 1 to 1708)

Komiyama, N.

Direct Submission

Submitted (02-APR-1994) to the DDBJ/EMBL/GenBank databases, Naoki

Komiyama, Meiji Institute of Health Science, 540 Naruda, Odawara,

Kanagawa 250, Japan (Tel:0465-37-3661, Fax:0465-36-2776)

2 (bases 1 to 1708)

Komiyama, N., Sone, T., Shimizu, K., Morikubo, K. and Kino, K.

CDNA cloning and expression of Cry j II, the second major allergen

of Japanese cedar pollen

Biochemical and Biophysical Research Communication 201, 1021-1028

(1994)

COMMENT Submitted (02-Apr-1994) to DDBJ by:

Naoki Komiyama

Meiji Institute of Health Science

540 Naruda, Odawara

Kanagawa 250

Japan

Phone: 0465-37-3661

Fax: 0465-36-2776.

Location/Qualifiers

1. 1708

/organism="Cryptomeria japonica"

/db_xref="taxon:3369"

45. 1589

/standard_name="Cry j II"

/function="second major allergen of Japanese cedar pollen"

/codon_start=1

/product="allergen"

/protein_id="BA06172.1"

/db_xref="PID:d1006734"

/db_xref="PID:9506858"

/db_xref="GI:506858"

/translation="MAMKLIAPMAFLAQLIIMAAEDSOAIMLDSVEKLRNSNS

LRKVERSHRDATNIFNVEKYGAVGDKDCDEASTAWACAKNPASALVPGSKFY

VNNLFRNGCPQPHFTFKYDGIATAIQNPASKNKNINWQPKLGTETLMGSGYIDGG

KOMWAGQCKWVNGREICNDNRPTAIKDFSTGLITQGLKLNSEPEHLYGKCEGK

IGISITAPRDSPTNIDGIDIFASKNFHLOKNTIGDQCVAILGTSSTIVIEDICG

GGISISIGSGRENSRAEVSIVVNGAKFIDIDNGLRITWGGSGMASHIIEYENMI

NSENPFLINQFCTGSASACONORSAVOIODVYKNIRGTSATAAAIQKCSDPCKD

IKLSDLSIKTSGKTAGCINCIONRANGFEGHVPACKNLSPAKRRKESSHKPTVMY

ENMRATDKGNRTRILLGSRPCTCKKCGSCCKAKLYIVIRIMQEYYPGRWICSCSH

GRIVHP"

BASE COUNT 568 a 294 c 380 g 466 t

alignment_scores:
 Quality: 99.00 Length: 19

Ratio: 5.211 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-142-524-1 x CPOCRXJII ..

Align seg 1/1 to: CPOCRXJII from: 1 to: 1708

31 IleaapllephalaSerlysaSnPhenHisLeugInlysaSnThIleGI 47
 |||||||
 750 ATTGATATCTTTCATCTAAACCTTCACTTACAAAGAACACGATAGG 799

47 yThrgly 49
 |||||||
 800 AACAGCG 806

seq_name: em_pat:E10716

seq_documentation_block:

ID E10716 standard: RNA; UNC: 1733 BP.

AC E10716:

SV E10716.1

XX d1109053

XX 08-OCT-1997 (Rel. 52, Created)

XX 08-OCT-1997 (Rel. 52, Last updated, Version 1)

XX cDNA encoding cedar pollen allergen, Cryj2.

XX JP 1996047392-A/1.

XX unidentified

XX unclassified.

XX [1]

XX 1-1733

XX Sone T., Komiyama N., Kii K.;

XX "ALLERGEN CRY J II EPI TOPE OF POLLEN OF JAPANESE CEDAR";

XX Patent number JP 1996047392-A/1, 20-FEB-1996.

XX MEIJI MILK PROD CO LTD.

XX OS Cryptomeria sp. (Cedar)

XX JP 1996047392-A/1

XX 20-FEB-1996

XX 07-NOV-1994 JP 1994297840

XX 05-NOV-1993 JP 93P 276773, 26-MAY-1994 JP 94P 134868

XX PC SOME TOSHIO, KOMIYAMA NAOKI, KII KOSUKE

XX C12N15/09,C07K7/08,C07K14/415//A61K39/36,C12Q1/68,G01N33/53;

XX CC strandness: Double;

XX CC topology: Linear;

XX CC hypothetical: No;

XX CC anti-sense: No;

XX CC Key

XX CC Location/Qualifiers

XX CC source

XX CC 1. 1733

XX CC /organism="Cryptomeria sp."

XX CC /tissue_type="Pollen"

XX CC 5'UTR

XX CC 45. 206

XX CC /product="signal peptide of cedar pollen

XX CC allergen, Cryj2"

XX CC 207. 1586

XX CC /product="cedar pollen allergen, Cryj2"

XX CC 45. 1589

XX CC /product="cedar pollen allergen, Cryj2"

XX CC 1590. 1733

XX CC 3'UTR

XX CC Location/Qualifiers

XX CC source

XX CC 1. 1733

XX CC /db_xref="taxon:32644"

XX CC /organism="unidentified"

alignment_scores: Quality: 99.00 Length: 19
 Ratio: 5.211 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-142-524-1 x E10716 ..

Align seg 1/1 to: E10716 from: 1 to: 1733

31 IleaapllephalaSerlysaSnPhenHisLeugInlysaSnThIleGI 47
 |||||||
 750 ATTGATATCTTTCATCTAAACCTTCACTTACAAAGAACACGATAGG 799

47 yThrgly 49
 |||||||
 800 AACAGCG 806

seq_name: gb_p11:CPOCJIP

seq_documentation_block:
 LOCUS CPOCJIP 1790 bp mRNA PIN 08-FEB-1999
 DEFINITION Cryptomeria japonica mRNA for Cry j II, complete cds.
 ACCESSION D37765
 NID 9377695
 VERSION D37765.1 GI:577695
 KEYWORDS Cry j II; allergen.
 SOURCE Cryptomeria japonica
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
 Taxodiaceae; Cryptomeria.
 1 (bases 1 to 1790)
 Nucleotide sequence
 Submitted (22-JUL-1994) to the DDBJ/EMBL/GenBank databases. Motoshi
 Namba, Fujisaki Institute, Hayashibara Biochemical
 Laboratories, Inc; Fujisaki 675-1, Okayama, Okayama 702, Japan
 (Tel:086-276-3141, Fax:086-276-6885)
 2 (bases 1 to 1790)
 Namba, M., Kurose, M., Torigoe, K., Hino, K., Taniguchi, Y., Fukuda, S.,
 Utsui, M. and Kurimoto, M.
 Molecular cloning of the second major allergen, Cry j II, from
 Japanese cedar pollen
 FEBS Lett. 353 (2), 124-128 (1994)
 95010777

JOURNAL MEDLINE
 FEATURES

location/Qualifiers
 1. 1790
 /organism="Cryptomeria japonica"
 /db_xref="taxon:3369"
 /tissue_type="pollen"
 126. 1670
 /codon_start=1
 /product="Cry j II precursor"
 /protein_id="BA07021.1"
 /db_xref="PID:d1007598"
 /db_xref="PID:g577696"
 /translation="MAMKFLAPAFVAMOLIIIAAEDQSAQIMLDSIDEOYLRNRS
 LRKVEHSRDAIINIFVEXKAVGDCGKHCTAFASTAWQAAKRSAMILVGNKKFV
 VNLFENGPCOPFTFKVGGITAAVONPASMKNNRIWLOFALITGLMGKVIDGOG
 KQWMAQCKWVNGREICNDPRPAIKEDFSTGLIIQGLKNSSEFLHVEGNCGVK
 IIGISTAPRDSPTDGDIDIFASKNFKLQNTIGTGDVVAIGTSSNIIYEDLIGP
 GHGISIGSLGRENSRAEVSIVYVNGAKFLDTONGRIKTMWGSGASNIIYENEMI
 NSENPILINOFCTISACONRSAYQIDQVYKNIRGNSATAAAIOLKCSNMECKD
 IRLSDISIKLTSGKISACINDNANGVFGSHVYPAKKNISPSAKRRESKSHRPKYVY
 KMGAYDKRNRIRILIGSRPPNCTNKCSCSCCKAKLVIRHINPOEYIPORWCSRH
 GKIIHP"

BASE COUNT 585 a 325 c 392 g 488 t
 ORIGIN

alignment_scores:

Quality: 99.00 Length: 19
Ratio: 5.211 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-142-524-1 x CPOCJIP ..

Align seg 1/1 to: CPOCJIP from: 1 to: 1790

31 IIEAEPHealaserLysAsnPhenHisLeuGlnLysAsnThrIleG1 47
|||||
831 ATGATATCTTGCATCTTAAACCTTCACTTACAAAGAACAGATAG 880

47 ythrlyg 49
|||||

881 AACAGCG 887

seq_name: em_pat:E11677

seq_documentation_block:
ID E11677 Standard; RNA; UNC; 1062 BP.

XX E11677;

XX E11677.1

XX d1110014

XX 08-OCT-1997 (Rel. 52, Created)

DT 08-OCT-1997 (Rel. 52, Last updated, Version 1)

XX cDNA encoding Cha o I.

KW JP 1996176192-A/1.

XX unidentified

XX unclassified.

XX [1]

RP 1-1062

RA Sone T., Komiyama N., Kii K.;

RT "ALLERGEN TO POLLEN OF CHAMAECYPARIS OBTUSA";

RL Patent number JP 1996176192-A/1, 09-JUL-1996.

RL MEITI MILK PROD CO LTD.

XX OS Chamaecyparis obtusa

CC PN JP 1996176192-A/1

CC PD 09-JUL-1996

CC PF 21-DEC-1994 JP 1994335089

CC PI SONE TOSHIO, KOMIYAMA NAOKI, KII KOSUKE

CC PC COTK14/415,C12N15/09,C12P21/02//A61K35/12,A61K35/64,A61K35/72,

CC PC A61K35/74,A61K39/36;

CC CC strandedness: Double;

CC CC topology: Linear;

CC FH key Location/Qualifiers

CC FT source 1..1062 /organism="Chamaecyparis obtusa"

CC FT /tissue_type="pollen"

CC FT mat_peptide 1..1062

CC FT /product="Cha o I"

XX key Location/Qualifiers

FH source 1..1062

FT /db_xref="taxon:32644"

FT /organism="unidentified"

XX SQ Sequence 1062 BP; 319 A; 198 C; 254 G; 291 T; 0 other;

alignment_scores:

Quality: 74.00 Length: 27
Ratio: 3.895 Gaps: 0
Percent Similarity: 70.370 Percent Identity: 59.259

alignment_block:

US-09-142-524-1 x E11677 ..

Align seg 1/1 to: E11677 from: 1 to: 1062

1 MetLysValThrValAlaPheAsnGlnPheGlyProAsnArgValPh 17
|||||
637 ATGAGGTGACAGTGGCATTCATTCATTTGGACCTAATGCTGCAACG 686

17 eileLysArgValSerAsnValIleIleHis 27
::: ||::: ::|||

687 AATGCCAGGCGCAGATATGACTATATACAT 717

seq_name: em_pat:E11678

seq_documentation_block:
ID E11678 Standard; RNA; UNC; 1125 BP.

XX E11678;

XX E11678.1

XX d1110015

XX 08-OCT-1997 (Rel. 52, Created)

DT 08-OCT-1997 (Rel. 52, Last updated, Version 1)

XX cDNA encoding Cha o I precursor.

KW JP 1996176192-A/2.

XX unidentified

XX unclassified.

XX [1]

RP 1-1125

RA Sone T., Komiyama N., Kii K.;

RT "ALLERGEN TO POLLEN OF CHAMAECYPARIS OBTUSA";

RL Patent number JP 1996176192-A/2, 09-JUL-1996.

RL MEITI MILK PROD CO LTD.

XX OS Chamaecyparis obtusa

CC PN JP 1996176192-A/2

CC PD 09-JUL-1996

CC PF 21-DEC-1994 JP 1994335089

CC PI SONE TOSHIO, KOMIYAMA NAOKI, KII KOSUKE

CC PC COTK14/415,C12N15/09,C12P21/02//A61K35/12,A61K35/64,A61K35/72,

CC PC A61K35/74,A61K39/36;

CC CC strandedness: Double;

CC CC topology: Linear;

CC FH key Location/Qualifiers

CC FT source 1..1125 /organism="Chamaecyparis obtusa"

CC FT /tissue_type="pollen"

CC FT sig_peptide 1..63

CC FT mat_peptide 64..1125

CC FT /product="Cha o I"

XX key Location/Qualifiers

FH source 1..1125

FT /db_xref="taxon:32644"

FT /organism="unidentified"

XX SQ Sequence 1125 BP; 327 A; 211 C; 264 G; 323 T; 0 other;

alignment_scores:

Quality: 74.00 Length: 27
 Ratio: 3.895 Gaps: 0
 Percent Similarity: 70.370 Percent Identity: 59.259

alignment_block:

US-09-142-524-1 x E11678 ..

Align seg 1/1 to: E11678 from: 1 to: 1125

1 MetLysValThrValAlaIaphaenGlnPheGlyProAsnArgArgValph 17
 |||
 700 ATGAAGGTGACAGTGCATTCATCATTTGGACCTAATGCTGACACACG 749

17 e11e1ysArGValSerAsnVal11e11e1His 27
 ::::: |||
 750 AATGCCAAGGCGACGATATGACTTATACAT 780

seq_name: gb_p11:D45404

seq_documentation_block: 1256 bp mRNA

LOCUS D45404 1256 bp mRNA PLN 10-FEB-1999
 DEFINITION Chamaecyparis obtusa male pollen mRNA for Chaol, complete cds.
 ACCESSION D45404
 KEYWORDS D45404.1 GI:1514942
 VERSION D45404.1
 SOURCE Chaol.
 ORGANISM Chamaecyparis obtusa male pollen cDNA to mRNA, clone: pCH-1.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
 Cupressaceae; Chamaecyparis.
 1 (bases 1 to 1256)
 Sone,T.
 Direct Submission
 Submitted (31-JAN-1995) to the DDBJ/EMBL/Genbank databases. Toshio
 Sone, Meiji Institute of Health Science, Department of
 Pharmaceutical Research; 540 Naruda, Odawara, Kanagawa 250, Japan
 (Tel:0465-37-3661, Fax:0465-36-2776)
 2 (bases 1 to 1256)
 Sone,T.
 Unpublished (1996)
 3 (sites)
 Suzuki,M., Komiyama,N., Itoh,M., Itoh,H., Sone,T., Kino,K.,
 Takagi,I., and Ohta,N.
 Purification, characterization and molecular cloning of Cha o 1, a
 major allergen of Chamaecyparis obtusa (Japanese cypress) pollen
 Mol. Immunol. 33 (4-5), 451-460 (1996)
 96265194

FEATURES

LOCATION/Qualifiers
 1..1256
 /organism="Chamaecyparis obtusa"
 /db_xref="taxon:13415"
 /clone="pCH-1"
 /sex="male"
 /tissue_type="pollen"
 1..1256
 /note="Chaol"
 50..1177
 /function="allergen"
 /codon_start=1
 /product="Chaol"
 /protein_id="BAA08246.1"
 /db_xref="PID:g1008841"
 /db_xref="PID:g1514943"
 /db_xref="GI:1514943"
 /translation="MASCITLAVLVFLCAIVSCFSDNPIDSCWRGDANWMDONRKLAD
 CAVFGSSAMGKGAFTYIVSSDDPVPNAPAGTLRYGATERSIMLIFSKNLKLIN
 MFLYLAGNKTIDRGAEVHIGNGPCLPMRTYSHVILHGLNHCNTSGNVLISEA
 SGVPVHADGDGAIITRMVNTDVWIDHNSLSDSDGLVDTVTLASTGVTISNHFENHKK
 VMLGHSIDIYSDKSMKVTVAENCFGNAGQRMPRARGLIHVANNNDPMSITAIIGG

sig_peptide 50..112
 mat_peptide 113..1174
 /product="Chaol"
 BASE COUNT 371 a 233 c 282 g 370 t
 ORIGIN

alignment_scores:

Quality: 74.00 Length: 27
 Ratio: 3.895 Gaps: 0
 Percent Similarity: 70.370 Percent Identity: 59.259

alignment_block:

US-09-142-524-1 x D45404 ..

Align seg 1/1 to: D45404 from: 1 to: 1256

1 MetLysValThrValAlaIaphaenGlnPheGlyProAsnArgArgValph 17
 |||
 749 ATGAAGGTGACAGTGCATTCATCATTTGGACCTAATGCTGACACACG 798

17 e11e1ysArGValSerAsnVal11e11e1His 27
 ::::: |||
 799 AATGCCAAGGCGACGATATGACTTATACAT 829

seq_name: em_pat:E11679

seq_documentation_block:

ID E11679 standard; RNA; UNC; 1260 BP.

XX E11679;

XX E11679.1

XX d1110016

XX 08-OCT-1997 (Rel. 52, Created)

DT 08-OCT-1997 (Rel. 52, Last updated, Version 1)

XX cDNA encoding Cha o 1.

XX JP 1996176192-A/3.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

XX unclassified.

CDS
38..496
/clone="pmc313-1"
/codon_start=1
/product="ribosomal protein S18"
/protein_id="CAA0750.1"
/db_xref="PID:9433447"
/db_xref="GI:433447"
/db_xref="SWISS-PROT:P25232"
/translation="MSLVPEKFOHILRVLTNTIDGRRIKIAFAITAKGVRRYAHV
LKRADIDLTRRAELETDEVERVITIMQNPQYKIPDWFNRQKDVKGKYSQVLANG
LDNKLREDLERLKRIRAHRGIRHFWGLRVGQHTKTGRGRRTGVSKK"

BASE COUNT 127 a 133 c 159 g 97 t
ORIGIN

alignment_scores:
Quality: 70.00 Length: 79
Ratio: 1.707 Gaps: 2
Percent Similarity: 51.899 Percent Identity: 30.380

alignment_block:
US-09-142-524-1 x RRRPS18A ..
Align seg 1/1 to: RRRPS18A from: 1 to: 516

```

12 ProAsnArgArgValPheIleLysArgValSerAsnValIleIleHisG1 28
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 CCCGAGAGTTTCAGACATCTCGAGTACTCAGACCAACATCGATGG 102
28 YArgArgIleAspIlePheIleAspSerLysAsnPheHisLeuGlnLysAsnT 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103 GCGGCGGAAATAGCCCTTCCTATCAGCTGCCATTAG..... 139
45 hrIleGlyThrGlyArgArg..... 51
|||||:|||||:|||||:|||||:|||||:|||||:
140 .....GGTGTGGGCGGAGATACGCTCATGTGTTGAGGAAGACAGAC 184
52 IleSerLeuLysLeuThrSerGlyLysIleAlaSerArgArgValAspG1 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
185 ATTGACCTCACCAGAGGCGCTGGGAGACTCAGCAGAGACGAGTGGACGC 234
68 YIleIleAlaAlaIleTyrGlnAsnProAlaSerTrpLys 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
235 TGTGATCACCATCATGACAGACCCAGACAAATACAAAG 271
seq_name: gb_ro:MUSKE3B

```

seq_documentation_block:
LOCUS MUSKE3B 525 bp mRNA ROD 16-MAR-1994
DEFINITION Mus musculus ribosomal protein (Ke-3) mRNA, complete cds.
ACCESSION M76763
NID 9198579
VERSION M76763.1 GI:198579
KEYWORDS ribosomal protein; ribosomal protein S13 homologue.
SOURCE Mus musculus cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 525)
AUTHORS MacMurray,A.J. and Shin,H.S.
TITLE The murine MHC encodes a mammalian homolog of bacterial ribosomal protein S13
JOURNAL Mamm. Genome 2 (2), 87-95 (1992)
MEDLINE 92182530
FEATURES
Location/Qualifiers
1..525
/organism="Mus musculus"
/db_xref="taxon:10090"
1..22
/partial
/note="putative"
23..525
/gene="Ke-3"
gene

CDS
23..481
/gene="Ke-3"
/note="putative"
/codon_start=1
/product="ribosomal protein"
/protein_id="AA16795.1"
/db_xref="PID:9198580"
/db_xref="GI:198580"
/translation="MSLVPEKFOHILRVLTNTIDGRRIKIAFAITAKGVRRYAHV
LKRADIDLTRRAELETDEVERVITIMQNPQYKIPDWFNRQKDVKGKYSQVLANG
LDNKLREDLERLKRIRAHRGIRHFWGLRVGQHTKTGRGRRTGVSKK"

3 UTR
polyA_signal
polyA_site
525
/gene="Ke-3"
/note="putative"
502..507
/gene="Ke-3"
/note="putative"
525
/gene="Ke-3"
/note="putative"

BASE COUNT 138 a 128 c 156 g 103 t
ORIGIN

alignment_scores:
Quality: 70.00 Length: 79
Ratio: 1.707 Gaps: 2
Percent Similarity: 51.899 Percent Identity: 30.380

alignment_block:
US-09-142-524-1 x MUSKE3B ..
Align seg 1/1 to: MUSKE3B from: 1 to: 525

```

12 ProAsnArgArgValPheIleLysArgValSerAsnValIleIleHisG1 28
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
38 CTGAGAGAGTTCAGACATTTTGGAGTACTCAGACCAACATCGATGG 87
28 YArgArgIleAspIlePheIleAspSerLysAsnPheHisLeuGlnLysAsnT 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 GCGGCGGAAATAGCCCTTCCTATCAGCTGCCATTAG..... 124
45 hrIleGlyThrGlyArgArg..... 51
|||||:|||||:|||||:|||||:|||||:|||||:
125 .....GCGTGGGCGGAGATACGCTCATGTGTTGAGGAAGACAGAC 169
52 IleSerLeuLysLeuThrSerGlyLysIleAlaSerArgArgValAspG1 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 ATGACCTCACCAGAGGCGCTGGGAGACTCAGCAGAGATGAGTGGAGCG 219
68 YIleIleAlaAlaIleTyrGlnAsnProAlaSerTrpLys 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 AGTGATCACCATCATGACAGACCCAGCAGACTACAG 256
seq_name: gb_pl2:AF106662

```

seq_documentation_block:
LOCUS AF106662 1104 bp mRNA PLN 10-JAN-1999
DEFINITION Juniperus ashei pollen allergen 1-1 mRNA, complete cds.
ACCESSION AF106662
NID 94138876
VERSION AF106662.1 GI:4138876
KEYWORDS
SOURCE Ozark white cedar.
ORGANISM Juniperus ashei
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
Cupressaceae; Juniperus.
REFERENCE 1 (bases 1 to 1104)
AUTHORS Midoro-Horiuchi,T.M., Goldblum,R.M., Kurosky,A., Wood,T.G. and Brooks,E.G.
TITLE Molecular cloning of mountain cedar (Juniperus ashei) pollen major allergen, Jun a 1

JOURNAL Unpublished
 2 (bases 1 to 1104)
 Mido-ro-Horiuti, T.M., Goldblum, R.M., Kurosky, A., Wood, T.G. and
 Brooks, E.G.
 TITLE Direct Submission
 Submitted (12-NOV-1998) Department of Pediatrics, Child Health
 Research Center, University of Texas Medical Branch, 301 University
 Blvd., Galveston, TX 77555-0366, USA
 FEATURES Location/Qualifiers
 source 1..1104
 /organism="Juniperus ashei"
 /db_xref="taxon:13101"
 1..1104
 /note="Jun a 1-1"
 CDS
 /codon_start=1
 /product="pollen major allergen 1-1"
 /protein_id="A003608.1"
 /db_xref="PID:94138879"
 /db_xref="GI:4138877"
 /translation="MASPCILAVIVLCAIVSCYSDNPIDSCMGSDSNMDONRMKLD
 CAVFGSSITMGKGGDEFTVYSTDNDVNPPTGLRIGATREKALMIFISQNMNITIK
 MPLYAGHKITIDGRGADVHLGNGPCLEFMKRVSHVLIHSIHGCVTSVLGDIVSES
 IGVEPVHODGDAITMRNVTNADIDHNSLSDSDGLIDVTLSGTITISNNHFNHKK
 VMLGHDPTVDDKSMKVTYVAFNFGPNAGORMPRARVGLVHANNNDPNNITAIIG
 SNNPTILSEGNSTAPSESYKKEVTKRIGCESPSACANWVWRSTDAFINGAVFVSSG
 KTEETNITNSNEAFKVENGNAAAPOLTKNAGVYI"
 BASE COUNT 324 a 202 c 264 g 314 t
 ORIGIN
 alignment_scores:
 Quality: 73.00 Length: 27
 Ratio: 3.842 Gaps: 0
 Percent Similarity: 70.370 Percent Identity: 55.556
 alignment_block:
 US-09-142-524-1 x AF106662 ..
 Align seg 1/1 to: AF106662 from: 1 to: 1104
 1 MetLysValThrValAlaIlePheAsnGlnPheGlyProAsnArgArgValAla 17
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 700 ATGAAAGTGACAGTGGCGTCAATTCATGACCTAATGCTGGCGCAAG 749
 17 eileLysArgValSerAsnValIleIleHis 27
 ::|||::|
 750 AATGCCAAGGCGACGATGACTTGATCAT 780
 seq_name: gb_p12:AF106663
 seq_documentation_block:
 LOCUS AF106663 1104 bp mRNA PLN 10-JAN-1999
 DEFINITION Juniperus ashei pollen major allergen 1-2 mRNA, complete cds.
 ACCESSION AF106663
 NID 94138878
 VERSION AF106663.1 GI:4138878
 KEYWORDS
 SOURCE Ozark white cedar.
 ORGANISM Juniperus ashei
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
 Cupressaceae; Juniperus.
 REFERENCE 1 (bases 1 to 1104)
 AUTHORS Mido-ro-Horiuti, T.M., Goldblum, R.M., Kurosky, A., Wood, T.G. and
 Brooks, E.G.
 TITLE Molecular cloning of mountain cedar (Juniperus ashei) pollen major
 allergen, Jun a 1
 JOURNAL Unpublished
 2 (bases 1 to 1104)
 Mido-ro-Horiuti, T.M., Goldblum, R.M., Kurosky, A., Wood, T.G. and
 Brooks, E.G.
 TITLE Direct Submission
 Submitted (12-NOV-1998) Department of Pediatrics, Child Health

Research Center, University of Texas Medical Branch, 301 University
 Blvd., Galveston, TX 77555-0366, USA
 FEATURES Location/Qualifiers
 source 1..1104
 /organism="Juniperus ashei"
 /db_xref="taxon:13101"
 1..1104
 /note="Jun a 1-2"
 CDS
 /codon_start=1
 /product="pollen major allergen 1-2"
 /protein_id="A003609.1"
 /db_xref="PID:94138879"
 /db_xref="GI:4138879"
 /translation="MASPCILAVIVLCAIVSCYSDNPIDSCMGSDSNMDONRMKLD
 CAVFGSSITMGKGGDEFTVYSTDNDVNPPTGLRIGATREKALMIFISQNMNITIK
 MPLYAGHKITIDGRGADVHLGNGPCLEFMKRVSHVLIHSIHGCVTSVLGDIVSES
 IGVEPVHODGDAITMRNVTNADIDHNSLSDSDGLIDVTLSGTITISNNHFNHKK
 VMLGHDPTVDDKSMKVTYVAFNFGPNAGORMPRARVGLVHANNNDPNNITAIIG
 SNNPTILSEGNSTAPSESYKKEVTKRIGCESPSACANWVWRSTDAFINGAVFVSSG
 KTEETNITNSNEAFKVENGNAAAPOLTKNAGVYI"
 BASE COUNT 324 a 203 c 264 g 313 t
 ORIGIN
 alignment_scores:
 Quality: 73.00 Length: 27
 Ratio: 3.842 Gaps: 0
 Percent Similarity: 70.370 Percent Identity: 55.556
 alignment_block:
 US-09-142-524-1 x AF106663 ..
 Align seg 1/1 to: AF106663 from: 1 to: 1104
 1 MetLysValThrValAlaIlePheAsnGlnPheGlyProAsnArgArgValAla 17
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 700 ATGAAAGTGACAGTGGCGTCAATTCATGACCTAATGCTGGCGCAAG 749
 17 eileLysArgValSerAsnValIleIleHis 27
 ::|||::|
 750 AATGCCAAGGCGACGATGACTTGATCAT 780
 seq_name: gb_com:AB000911
 seq_documentation_block:
 LOCUS AB000911 528 bp mRNA MAM 05-FEB-1999
 DEFINITION Sus scrofa mRNA for ribosomal protein, complete cds.
 ACCESSION AB000911
 NID 91841303
 VERSION AB000911.1 GI:1841303
 KEYWORDS ribosomal protein S13.
 SOURCE Sus scrofa CDNA to mRNA.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Artiodactyla; Suidae; Sus.
 REFERENCE 1 (bases 1 to 528)
 AUTHORS Hamasima, N.
 TITLE Direct Submission
 JOURNAL Submitted (07-FEB-1997) to the DDBJ/EMBL/GenBank databases.
 Naitiyuki Hamasima, STAF-Institute, Animal Genome Research Program
 Team; 446-1 Ippaiyuka, Kamiyokoda, Tsukuba, Ibaraki 305, Japan
 (E-mail:hamasima@genome.staif.or.jp, Tel:0298-38-2190,
 Fax:0298-38-2337)
 2 (sites)
 AUTHORS Kimura, M., Kawakami, K., Suzuki, H. and Hamasima, N.
 TITLE Cloning of the pig homolog of bacterial ribosomal protein S13
 JOURNAL Unpublished (1997)
 FEATURES Location/Qualifiers
 source 1..528
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 30..488
 /codon_start=1

```

/product="ribosomal protein"
/protein_id="BAA19211.1"
/db_xref="PID:d1019961"
/db_xref="PID:q1841304"
/db_xref="GI:1841304"
/translation="MSLVPEKFOHILRVLTNIDGRKIAFAITATKGYGRYAHY
LKKADIDLRKAGELEVEDEVERITIMONPROYKIPDKFLNQKDYADGKYSOVLANG
LDNKLREDELERLKIIRAHGRLRHEWGLEVRGQHKITGRGRIVGVSKK"
BASE COUNT      144 a      127 c      152 g      105 t
ORIGIN

```

```

alignment_scores:
  quality: 70.00      length: 79
  ratio: 1.707      gaps: 2
  percent similarity: 51.899      percent identity: 30.380

```

```

alignment block:
US-09-142-524-1 x AB000911 ..

```

```

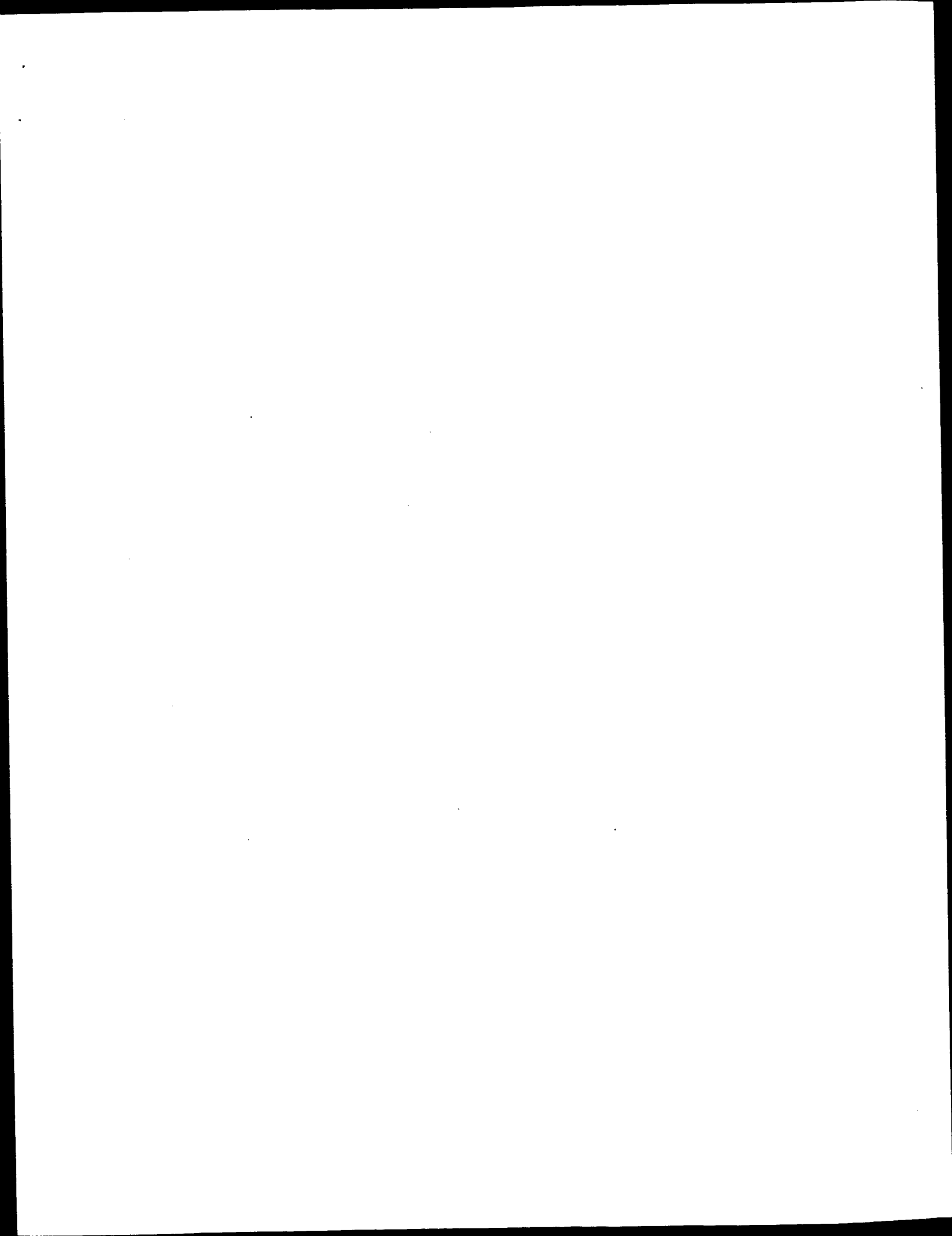
Align seq 1/1 to: AB000911 from: 1 to: 528

```

```

12 ProAsnArgArgValPheIleIleLysArgValSerAsnValIleIleHisG1 28
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
45 CCGAGAGAGTTCACGACACATTTCGAGTACTCAACACACACATCGATGG 94
28 YArgArgIleAspIlePheAlaSerLysAsnPheHisLeuGlnLysAsnT 45
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
95 GCGGCGGAAATATGCTTGCCTATCATCTGCATTAAG..... 131
45 hrlleGlyThrGlyArg..... 51
|||:::|||||:::|||||:::|||||:::|||||:::|||||
132 ....GCTGAGGACGAGATATGCTATGCTGTGAGAAACAGAC 176
52 lIleSerLeuLysLeuThrSerGlyLysIleAlaSerArgArgValAspI 68
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
177 ATGACCTCACCACAGAGGCGACAGAGCTCACAGAGATGAGGTGGAACG 226
68 yIleIleAlaAlaTyGlnAsnProAlaSerTrpLys 80
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
227 TGTAATCACCATATATGACAGATCTCGCCAAATACAAAG 263

```



OM of: US-09-142-524-1 to: N_Geneseq_36.* out_format : pfs

Date: Sep 29, 1999 2:01 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-1998 Comugen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xip
- /s2n2.1/USPRO.spool/US09142524/rnatc.29091999.125959.12792/APP-query.fasta.1
-N_Geneseq_36 -QWMT-fastap -SUFFIX=ring -GAPOP=12.000
-GAPOP=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=6.000 -GAPOP=10.000 -XGAPOP=10.000 -XGAPOP=10.000
-DELOP=6.000 -DELOP=7.000 -XGAPOP=10.000 -XGAPOP=10.000
-TRANS=human40.cdi -LIST=45 -MATRIX=blissum62
-ALIGN=15 -MODE=LOCAL -OUTWMT=pfs -NORM=stat -USER=US09142524
-NCPU=6 -ICPU=3 -WAIT -THREADS=1

Search information block:

Query: US-09-142-524-1
Query length: 80
Database: N_Geneseq_36.*
Database sequences: 311585
Database length: 125096042
Search time (sec): 115.290000

Score list:

Sequence	Strd Ord	ZScore	EScore	Len	Documentation
N_Geneseq_36:Q84044	+	99.00	235.28	176-05	1380
N_Geneseq_36:Q86050	+	99.00	235.16	176-05	1395
N_Geneseq_36:Q86051	+	99.00	235.04	176-05	1410
N_Geneseq_36:Q86052	+	99.00	234.51	176-05	1429
N_Geneseq_36:Q86053	+	99.00	234.05	2.0e-05	1542
N_Geneseq_36:Q86054	+	99.00	232.81	2.3e-05	1733
N_Geneseq_36:Q86055	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86056	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86057	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86058	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86059	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86060	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86061	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86062	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86063	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86064	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86065	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86066	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86067	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86068	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86069	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86070	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86071	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86072	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86073	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86074	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86075	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86076	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86077	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86078	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86079	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86080	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86081	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86082	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86083	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86084	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86085	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86086	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86087	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86088	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86089	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86090	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86091	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86092	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86093	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86094	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86095	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86096	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86097	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86098	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86099	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86100	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86101	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86102	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86103	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86104	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86105	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86106	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86107	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86108	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86109	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86110	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86111	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86112	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86113	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86114	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86115	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86116	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86117	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86118	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86119	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86120	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86121	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86122	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86123	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86124	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86125	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86126	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86127	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86128	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86129	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86130	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86131	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86132	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86133	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86134	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86135	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86136	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86137	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86138	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86139	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86140	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86141	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86142	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86143	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86144	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86145	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86146	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86147	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86148	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86149	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86150	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86151	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86152	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86153	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86154	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86155	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86156	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86157	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86158	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86159	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86160	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86161	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86162	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86163	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86164	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86165	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86166	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86167	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86168	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86169	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86170	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86171	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86172	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86173	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86174	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86175	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86176	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86177	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86178	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86179	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86180	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86181	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86182	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86183	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86184	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86185	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86186	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86187	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86188	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86189	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86190	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86191	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86192	+	99.00	232.76	2.3e-05	1733
N_Geneseq_36:Q86193</					

PT Allergic Cry j II protein and fragments from Japanese cedar
 TT pollen - used to diagnose, treat and prevent Japanese cedar
 PT pollinosis
 PS Disclosure: Page 49; 89pp; English.
 CC The sequence is of a Japanese cedar pollen allergen Cry j IIA
 CC fragment. The DNA and its fragments can be used for diagnosis and
 CC treatment of Japanese cedar pollinosis and to identify similar
 CC sequences in other plants.
 CC See also Q66048-67.
 SQ Sequence 1395 BP; 460 A; 257 C; 318 G; 360 T;

alignment_scores:
 Quality: 99.00 Length: 19
 Ratio: 5.211 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-142-524-1 x Q66050 ..

Align seg 1/1 to: Q66050 from: 1 to: 1395

31 ILEASPILEPHEALASERLYSASNPHEHISLEUGLNLYSASNTHTLLEGI 47
 |||||||
 556 ATTGATATCTTTCATCTAAACCTTCACTTACAAAGAACACGATAGG 605
 47 yThrgly 49
 |||||||
 606 AACAGGG 612

seq_name: N_Geneseq_36:Q66049

seq_documentation_block:

ID Q66049 standard; DNA: 1410 BP.
 AC Q66049;
 DT 01-FEB-1995 (first entry)
 DE Japanese cedar pollen allergen Cry j II DNA fragment.
 KM Cedar pollinosis; diagnostic; ss.
 OS Crytomeria japonica.
 PN WO9411512-A.
 PD 26-MAY-1994.
 PE 12-NOV-1993; U11000.
 PR 12-NOV-1992; US-975179.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Brauer A, Kuo M, Pollock J, Yeung S;
 DR WPI: 94-183513/22.
 PT Allergic Cry j II protein and fragments from Japanese cedar
 PT pollen - used to diagnose, treat and prevent Japanese cedar
 PT pollinosis
 PS Disclosure: Page 48; 89pp; English.
 CC The sequence is of a Japanese cedar pollen allergen Cry j IIA
 CC fragment. The DNA and its fragments can be used for diagnosis and
 CC treatment of Japanese cedar pollinosis and to identify similar
 CC sequences in other plants.
 CC See also Q66048-67.
 SQ Sequence 1410 BP; 467 A; 258 C; 322 G; 363 T;

alignment_scores:
 Quality: 99.00 Length: 19
 Ratio: 5.211 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-142-524-1 x Q66049 ..

Align seg 1/1 to: Q66049 from: 1 to: 1410

31 ILEASPILEPHEALASERLYSASNPHEHISLEUGLNLYSASNTHTLLEGI 47
 |||||||
 571 ATTGATATCTTTCATCTAAACCTTCACTTACAAAGAACACGATAGG 620
 47 yThrgly 49

||||||
 621 AACAGGG 627

seq_name: N_Geneseq_36:Q66051

seq_documentation_block:

ID Q66051 standard; DNA: 1479 BP.
 AC Q66051;
 DT 01-FEB-1995 (first entry)
 DE Japanese cedar pollen allergen Cry j II DNA fragment.
 KM Cedar pollinosis; diagnostic; ss.
 OS Crytomeria japonica.
 PN WO9411512-A.
 PD 26-MAY-1994.
 PE 12-NOV-1993; U11000.
 PR 12-NOV-1992; US-975179.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Brauer A, Kuo M, Pollock J, Yeung S;
 DR WPI: 94-183513/22.
 PT Allergic Cry j II protein and fragments from Japanese cedar
 PT pollen - used to diagnose, treat and prevent Japanese cedar
 PT pollinosis
 PS Disclosure: Page 50; 89pp; English.
 CC The sequence is of a Japanese cedar pollen allergen Cry j IIA
 CC fragment. The DNA and its fragments can be used for diagnosis and
 CC treatment of Japanese cedar pollinosis and to identify similar
 CC sequences in other plants.
 CC See also Q66048-67.
 SQ Sequence 1479 BP; 490 A; 268 C; 337 G; 384 T;

alignment_scores:
 Quality: 99.00 Length: 19
 Ratio: 5.211 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-142-524-1 x Q66051 ..

Align seg 1/1 to: Q66051 from: 1 to: 1479

31 ILEASPILEPHEALASERLYSASNPHEHISLEUGLNLYSASNTHTLLEGI 47
 |||||||
 640 ATTGATATCTTTCATCTAAACCTTCACTTACAAAGAACACGATAGG 689
 47 yThrgly 49
 |||||||
 690 AACAGGG 696

seq_name: N_Geneseq_36:Q84045

seq_documentation_block:

ID Q84045 standard; cDNA: 1542 BP.
 AC Q84045;
 DT 27-SEP-1995 (first entry)
 DE Japonicum allergen cDNA.
 KM Japonicum allergen; induced histamine release; antiallergic peptide;
 KW IgE cross-linking inhibition; ds.
 OS Japonicum sp.
 FH Key Location/Qualifiers
 FT mat_peptide 1..1542
 FT /*tag= a
 PN WO9502412-A.
 PD 26-JAN-1995.
 PE 15-JUL-1994; J01164.
 PR 16-JUL-1993; JP-177008.
 PR 01-SEP-1993; JP-217725.
 PR 07-APR-1994; JP-069336.
 PA (MEIP) MEIJI MILK PROD CO LTD.
 PI Kino K, Kohno Y, Komiyama N, Sone T;
 DR WPI: 95-067159/09.
 DR P-PDB: R69792.
 PT Peptide antiallergic agent - inhibits cross-linking of allergen

PT with 19E antibody
 PS Disclousure; Pages 27-28; 46pp: Japanese.
 CC Q84045 encodes R69792 Japonicum allergen, from which the antiallergic
 CC peptides R69845-R69809 were derived. The peptides ability to inhibit
 CC the cross-linking of an allergen, to an 19E antibody can be used in
 SQ the prevention and treatment of allergic diseases.
 SQ Sequence 1542 BP; 506 A; 280 C; 352 G; 404 T;

alignment_scores:

Quality: 99.00 Length: 19
 Ratio: 5.211 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-142-524-1 x Q84045 ..

Align seg 1/1 to: Q84045 from: 1 to: 1542

31 IleaspllephelaSerlysaSnphenisIeugInlysaSnhrlllegl 47
 |||||||
 706 ATTGATATCTTGCACTTAATAACTTTCACCTACAAAGAACAGATAGG 755
 47 yThrcly 49
 |||||||
 756 AACACGGG 762

seq_name: N_Geneseq_36:Q90156

seq_documentation_block:

ID Q90156 standard; DNA; 1545 BP.
 AC Q90156;
 DT 01-NOV-1995 (first entry)
 DE Japanese cedar pollen allergen gene.
 KW Japanese cedar; pollen; allergen; allergy; therapy; diagnostic;
 KM desensitizer; ss
 OS Cryptomeria japonica
 FH Key Location/Qualifiers
 FT cds 1..1545
 FT EP-655500-A.
 PD 31-MAY-1995.
 PE 03-NOV-1994; 308117.
 PR 05-NOV-1993; JP-299151.
 PR 20-DEC-1993; JP-344596.
 PR 27-DEC-1993; JP-346814.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PI Kurimoto M, Namba M, Torigoe K;
 DR P-PSDB; R74333.
 PT New Japanese cedar pollen allergen polypeptide - and DNA coding for it,
 PT useful for treatment and diagnosis of cedar pollen allergy
 PS Claim 6; Page 29; 41pp; English.
 CC The gene encoding an allergen of Japanese cedar pollen was isolated
 CC by PCR amplification using primers based on portions of the allergen
 CC protein. The gene was used for recombinant allergen production in
 CC E. coli (vector plasmid pK-223-3).
 SQ Sequence 1545 BP; 508 A; 283 C; 351 G; 403 T;

alignment_scores:

Quality: 99.00 Length: 19
 Ratio: 5.211 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-142-524-1 x Q90156 ..

Align seg 1/1 to: Q90156 from: 1 to: 1545

31 IleaspllephelaSerlysaSnphenisIeugInlysaSnhrlllegl 47
 |||||||
 706 ATTGATATCTTGCACTTAATAACTTTCACCTACAAAGAACAGATAGG 755

47 yThrcly 49
 |||||||
 756 AACACGGG 762

seq_name: N_Geneseq_36:Q66048

seq_documentation_block:

ID Q66048 standard; CDNA; 1726 BP.
 AC Q66048;
 DT 01-FEB-1995 (first entry)
 DE Japanese cedar pollen allergen Cry j II.
 KW Cedar pollinosis; diagnostic; ss.
 OS Cryptomeria japonica.
 FH Key Location/Qualifiers
 FT cds 42..1586
 FT MO9411512-A.
 PD 26-MAY-1994.
 PE 12-NOV-1993; U11000.
 PR 12-NOV-1992; US-975179.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Brauer A, Kuo M, Pollock J, Yeung S;
 DR WPI; 94-183513/22.
 DR P-PSDB; R53690.
 PT Allergenic Cry j II protein and fragments from Japanese cedar
 PT pollen - used to diagnose, treat and prevent Japanese cedar
 PT pollinosis
 PS Claim 2; Fig 4; 89pp; English.
 CC The sequence is that encoding a Japanese cedar pollen allergen Cry j
 CC II. The DNA and its fragments can be used for diagnosis and
 CC treatment of Japanese cedar pollinosis and to identify similar
 CC sequences in other plants.
 SQ See also Q66049-67.
 SQ Sequence 1726 BP; 579 A; 299 C; 383 G; 465 T;

alignment_scores:

Quality: 99.00 Length: 19
 Ratio: 5.211 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-142-524-1 x Q66048 ..

Align seg 1/1 to: Q66048 from: 1 to: 1726

31 IleaspllephelaSerlysaSnphenisIeugInlysaSnhrlllegl 47
 |||||||
 747 ATTGATATCTTGCACTTAATAACTTTCACCTACAAAGAACAGATAGG 796
 47 yThrcly 49
 |||||||
 797 AACACGGG 803

seq_name: N_Geneseq_36:Q84046

seq_documentation_block:

ID Q84046 standard; CDNA; 1733 BP.
 AC Q84046;
 DT 27-SEP-1995 (first entry)
 DE Japonicum allergen cDNA and flanking sequences.
 KW Japonicum allergen; induced histamine release; antiallergic peptide;
 KW 19E cross-linking inhibition; flanking sequences; ds.
 OS Japonicum sp.
 FH Key Location/Qualifiers
 FT cds 45..1589
 FT MO9502412-A.
 PD 26-JAN-1995.
 PE 15-JUL-1994; J01164.
 PR 16-JUL-1993; JP-177008.
 PR 01-SEP-1993; JP-217725.

PR 07-APR-1994; JP-069336.
 PA (MEIP) MEIJI MILK PROD CO LTD.
 PI KINO K, Kohno Y, Komiyama N, Sone T;
 DR WPI; 95-067159/09.
 DR P-PSDB; R69792.
 PT Peptide antiallergic agent - inhibits cross-linking of allergen
 with IgE antibody
 PS Example 1: Page 29; 46pp; Japanese.
 CC 084046 encodes R69792 Japonicum allergen, from which the antiallergic
 peptides R69845-R69809 were derived. The peptides ability to inhibit
 the cross-linking of an allergen, to an IgE antibody can be used in
 the prevention and treatment of allergic diseases.
 CC Sequence 1733 BP; 593 A; 294 C; 380 G; 466 T;
 SQ

alignment_scores:
 Quality: 99.00 Length: 19
 Ratio: 5.211 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-142-524-1 x 084046 ..

Align seg 1/1 to: 084046 from: 1 to: 1733

31 ILeAspIlePheAlaSerLysAsnPhenHisLeuGlnLysAsnThrIleGI 47
 |||||||
 750 ATGTGATATCTTTGCATCTAAACCTTCACTACAAAGAACAGCATAGG 799

47 yThrGly 49
 |||||||
 800 AACACGGG 806

seq_name: N_Geneseq_36:T18102

seq_documentation_block:

ID T18102 standard; cDNA to mRNA; 1733 BP.

AC T18102:

DT 16-AUG-1996 (first entry)

DE Japan cedar pollen cDNA encoding Cry j II allergen.

KW Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;

KM Sugl pollinosis; diagnosis; treatment; ss.

OS Cryptomeria japonica.

PT Key Location/Qualifiers

FT 5'utr 1..44

FT cds /tag= a

FT signal_peptide 45..206

FT mat_peptide 207..1586

FT 3'utr 1590..1733

FT /tag= d

FT /tag= e

PN J08047392-A.

DT 20-FEB-1996.

DT 07-NOV-1994; 297840.

DT 05-NOV-1993; JP-276773.

DT 26-MAY-1994; JP-134868.

PA (MEIP) MEIJI MILK PROD CO LTD.

DR WPI; 96-166249/17.

DR P-PSDB; R93599.

PT Japan cedar pollen allergen Cry j II epitope - comprises at least

part of specified 460 amino acid protein

PS Claim 2: Page 13-14; 17pp; Japanese.

CC T18102 encodes a Japan cedar pollen Cry j II allergen which

is useful in the diagnosis, prevention and treatment of Sugl

pollinosis, the allergic reaction to Japan cedar pollen. Significant

regions of the allergen were identified using overlapping peptides

of the full epitope derived from a Cry j II antigen specific T cell

line (see R97871-R97960). Amino acids 66-80 (R97884) and 186-200

CC (R978908) of the full mature 460 amino acid allergen are the most

SQ Sequence 1733 BP; 593 A; 294 C; 380 G; 466 T;

alignment_scores:
 Quality: 99.00 Length: 19
 Ratio: 5.211 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-142-524-1 x T18102 ..

Align seg 1/1 to: T18102 from: 1 to: 1733

31 ILeAspIlePheAlaSerLysAsnPhenHisLeuGlnLysAsnThrIleGI 47
 |||||||
 750 ATGTGATATCTTTGCATCTAAACCTTCACTACAAAGAACAGCATAGG 799

47 yThrGly 49
 |||||||
 800 AACACGGG 806

seq_name: N_Geneseq_36:T38518

seq_documentation_block:

ID T38518 standard; cDNA; 1125 BP.

AC T38518:

DT 28-NOV-1996 (first entry)

DE Chamaecyparis obtusa pollen allergen Cha o I cDNA (B).

KW Pollen allergen; Cha o I; T-cell epitope; prevention; treatment;

KM Pollinosis; ds.

OS Chamaecyparis obtusa.

PT Key Location/Qualifiers

FT cds 1..1125

FT /tag= a

FT /note= "STOP codon absent"

PN J08176192-A.

DT 09-JUL-1996.

DT 21-DEC-1994; 335089.

DT 21-DEC-1994; JP-335089.

PA (MEIP) MEIJI MILK PROD CO LTD.

DR WPI; 96-368225/37.

DR P-PSDB; W04345.

PT DNA encoding chamaecyparis obtusa pollen allergen - T cell

epitope(s) of which are useful in development of preventative and

treating agent for C. obtusa pollen pollinosis

PS Claim 10; Page 13; 17pp; Japanese.

CC The present sequence encodes the C. obtusa pollen allergen Cha o I,

the T-cell epitopes of which can be used in the development of a

preventive and treating agent for C. obtusa pollen pollinosis.

CC C. obtusa pollen (2.4 kg) was degreased with diethyl ether, and

dried at room temp. overnight. Cha o I was sepd. from it and

purified. RNA was extracted from C. obtusa pollen, and mRNA and

cDNA derived.

CC Sequence 1125 BP; 327 A; 211 C; 264 G; 323 T;

SQ

alignment_scores:
 Quality: 74.00 Length: 27
 Ratio: 3.895 Gaps: 0
 Percent Similarity: 70.370 Percent Identity: 59.259

alignment_block:
 US-09-142-524-1 x T38518 ..

Align seg 1/1 to: T38518 from: 1 to: 1125

1 MetLysValThrValAlaPheAsnGlnPheGlyProAsnArgValPhe 17
 |||||||
 700 ATGAAGGTGACAGTGCATTCATCAATTTGACCTAATGCTGCAGCAACG 749
 17 eLLeLysArgValLysAsnValIleIleHis 27
 ::||::|
 750 AATGCCAAGGGCAGCATATGACCTATATCAT 780

```

seq_name: N_Geneseq_36:T38519
seq_documentation_block:
ID T38519 standard; cDNA; 1260 BP.
AC T38519;
DE 28-NOV-1996 (first entry)
DE Chamaecyparis obtusa pollen allergen Cha o I cDNA (A).
KW Pollen allergen; Cha o I; T-cell epitope; prevention; treatment;
OS Chamaecyparis obtusa.
FH Key Location/Qualifiers
FT cds 113..1177
FT /tag= a
FT /note= "sequence comprising nucleotides 113..1174
PN J08176192-A.
PD 09-JUL-1996.
PF 21-DEC-1994; 335089.
PR 21-DEC-1994; JP-335089.
RA (MEIP) MEIJI MILK PROD CO LTD.
DR WPI; 96-368225/37.
DR P-PSDB; W04344.
DR DNA encoding chamaecyparis obtusa pollen allergen - T cell
epitope(s) of which are useful in development of preventative and
PT treating agent for C. obtusa pollen pollinosis
PS Claim 4; Pages 13-14; 17pp; Japanese.
CC The present sequence encodes the C. obtusa pollen allergen Cha o I,
CC the T-cell epitopes of which can be used in the development of a
CC preventive and treating agent for C. obtusa pollen pollinosis.
CC C. obtusa pollen (2.4 kg) was degassed with diethyl ether, and
CC dried at room temp. overnight. Cha o I was sepd. from it and
CC purified. RNA was extracted from C. obtusa pollen, and mRNA and
CC cDNA derived.
SQ Sequence 1260 BP; 375 A; 233 C; 282 G; 370 T;

alignment_scores:
Quality: 74.00 Length: 27
Ratio: 3.895 Gaps: 0
Percent Similarity: 70.370 Percent Identity: 59.259

alignment_block:
US-09-142-524-1 x T38519 ..
Align seg 1/1 to: T38519 from: 1 to: 1260

1 MetLysValThrValAlaIapheAnGlnPheGlyProAsnArgValph 17
|||||
749 ATGAAGTACACAGTGCATTCATCAATTGGACCTAATGCTGACACAG 798
|||
17 eileLysArgValSerAsnValIleIleHis 27
::: |||||
799 AATGCCAAGGCACGATATGACTATATCAT 829

seq_name: N_Geneseq_36:Q55272
seq_documentation_block:
ID Q55272 standard; cDNA; 1170 BP.
AC Q55272;
DE 13-JUL-1994 (first entry)
DE Jun s I gene.
KW Japanese cedar; pollen allergen; allergy; treatment; diagnosis;
KW T cell epitope; sensitivity; detection; ss.
OS Juniperus sabinooides.
FH Key Location/Qualifiers
FT cds 26..1129
FT /tag= a
FT signal_peptide 26..88
FT /tag= b
FT mat_peptide 89..1129
FT /tag= c
PN W09401560-A.

```

```

PD 20-JAN-1994.
PF 15-JAN-1993; U00139.
PR 10-JUL-1992; WC-U05661.
PR 01-SEP-1992; US-938990.
RA (IMMO-) IMMUNOLOGIC PHARM CORP.
PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J;
DR WPI; 94-035066/04.
DR P-PSDB; R45577.
DR Antigens derived from Japanese cedar pollen allergen Cry j I -
PT contain at least two T cell epitope(s), used to treat or diagnose
PT allergy
PS Disclosure; Fig 16; 137pp; English.
CC The sequence is that encoding Jun s I, a homologue of the Japanese
CC cedar pollen allergen Cry j I. Antigenic peptides derived from it
CC can be used for the treatment and diagnosis of allergies associated
CC with Japanese cedar pollen.
SQ Sequence 1170 BP; 350 A; 215 C; 268 G; 337 T;

alignment_scores:
Quality: 73.00 Length: 27
Ratio: 3.842 Gaps: 0
Percent Similarity: 70.370 Percent Identity: 55.556

alignment_block:
US-09-142-524-1 x Q55272 ..
Align seg 1/1 to: Q55272 from: 1 to: 1170

1 MetLysValThrValAlaIapheAnGlnPheGlyProAsnArgValph 17
|||||
725 ATGAAGTACACAGTGCCTCATCAATTGGACCTAATGCTGCGCAAG 774
|||
17 eileLysArgValSerAsnValIleIleHis 27
::: |||||
775 AATGCCAAGGCACGATATGACTTGTACAT 805

seq_name: N_Geneseq_36:T38521
seq_documentation_block:
ID T38521 standard; cDNA; 1772 BP.
AC T38521;
DE 28-NOV-1996 (first entry)
DE Chamaecyparis obtusa pollen allergen Cha o II cDNA.
KW Pollen allergen; Cha o II; T-cell epitope; prevention; treatment;
OS Chamaecyparis obtusa.
FH Key Location/Qualifiers
FT cds 32..1576
FT /tag= a
FT /note= "sequence comprising nucleotides 32..1573
PN J08176192-A.
PD 09-JUL-1996.
PF 21-DEC-1994; 335089.
PR 21-DEC-1994; JP-335089.
RA (MEIP) MEIJI MILK PROD CO LTD.
DR WPI; 96-368225/37.
DR P-PSDB; W04346.
DR DNA encoding chamaecyparis obtusa pollen allergen - T cell
PT epitope(s) of which are useful in development of preventative and
PT treating agent for C. obtusa pollen pollinosis
PS Claim 16; Page 16; 17pp; Japanese.
CC The present sequence encodes the C. obtusa pollen allergen
CC Cha o II, the T-cell epitopes of which can be used in the
CC development of a preventative and treating agent for C. obtusa
CC pollen pollinosis; C. obtusa pollen (2.4 kg) was degassed with
CC diethyl ether, and dried at room temp. overnight. Cha o II was
CC sepd. from it and purified. RNA was extracted from C. obtusa
CC pollen, and mRNA and cDNA derived.
SQ Sequence 1772 BP; 581 A; 328 C; 388 G; 475 T;

```

alignment_scores: length: 19
 Quality: 74.00
 Ratio: 4.111
 Percent Similarity: 94.737 Percent Identity: 73.684

alignment_block:
 US-09-142-524-1 x T38521 ..

Align seg 1/1 to: T38521 from: 1 to: 1772

```

31 1leasplephea1aserlyasnphehisreuglnlysasnhrilegl 47
   |||||||.....:|||||:|||||:|||||:|||||:|||||
737 ATTGATATCTTTCATCTTAAGAATTTCACATAGAAAGTCGTAATAGG 786
47 yThrGly 49
   |||||
787 AACAGGG 793

```

seq_name: N_Geneseq_36:071601

seq_documentation_block:
 ID Q71601 standard; cDNA to mRNA; 1317 BP.

AC Q71601;
 NT 24-MAR-1995 (first entry)
 Japanese cedar pollen antigen Cryj1 cDNA.
 Japanese cedar pollen antigen; allergen; Cryj1; sugi; pollinosis;

US Cryptomeria japonica.
 Location/Qualifiers

```

..      signal_peptide 62..124      /*tag= a
FT      mat_peptide 125..1183      /*tag= b
FT      3'utr 1184..1317      /*tag= c
FT      /*tag= d

```

PN J06197768-A.
 PD 19-JUL-1994.
 PR 07-JAN-1993; JP-001116.
 PA (METU) MEIJI SEIKA KAISHA.
 DR WPI: 94-268680/33.
 DR P-PSDB: R60166.

PT Sugi (Japanese cedar) pollen antigen Cryj1 - is useful for diagnosis, treatment and prevention of sugi pollinosis

PS Claim 6: Page 5-7; 9pp: Japanese.
 CC The coding sequence for the Japanese cedar ("sugi") pollen allergen Cryj1 was isolated from a cDNA library prepared from polyA mRNA. All CC or part of the Cryj1 protein can be used for diagnosis, treatment CC and prevention of sugi pollinosis.
 SQ Sequence 1317 BP; 410 A; 230 C; 284 G; 393 T;

alignment_scores: length: 27
 Quality: 71.00
 Ratio: 3.737
 Percent Similarity: 70.370 Percent Identity: 55.556

alignment_block:
 US-09-142-524-1 x Q71601 ..

Align seg 1/1 to: Q71601 from: 1 to: 1317

```

1 MetLysValThrValAlaIleAsnGlnPheGlyProAsnArgValPh 17
   |||||||.....:|||||:|||||:|||||:|||||:|||||
761 ATGAAGGTGACAGTGGCGCTTCAATCATTTGACCTAAGTGGACAAAG 810
17 eileLysArgValSerAsnValIleIleHis 27
   ::|||::|
811 AATGCCAGGCGACGATATGACTGTACAT 841

```

seq_name: N_Geneseq_36:Q35304

seq_documentation_block:

ID Q35304 standard; cDNA to mRNA; 1337 BP.

AC Q35304;
 NT 03-JUN-1993 (first entry)
 DE Cry j 1 gene.
 KW Japanese cedar pollen; allergen; antigen; allergy; B cell; T cell; ss.
 OS Cryptomeria japonica.
 FH Key location/Qualifiers

```

FT      cds 66..1187
FT      signal_peptide 66..128      /*tag= a
FT      /*tag= b
FT      mat_peptide 129..1187      /*tag= c

```

PN M09301213-A.
 PD 21-JAN-1993.
 PR 10-JUL-1992; U05661.
 PR 12-JUL-1991; US-728134.
 PR 15-JUL-1991; US-730452.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Griffith J, Pollock J;
 DR WPI: 93-045434/05.
 DR P-PSDB: R31937.

PT Nucleic acid sequence encoding Cryptomeria japonica allergen - for the diagnosis treatment and prevention of allergic reactions

PT to Japanese cedar pollen
 PS Claim 1; Page 40; 69pp; English.
 CC Fresh pollen and staminate cone samples were collected from a single CC Cryptomeria japonica (Japanese cedar) tree. RNA was prep'd. and used CC to synthesise cDNA. The cDNA was subjected to successive rounds of CC PCR to yield a 193 bp clone, JCT1.6, found to contain a partial CC sequence of Cry j 1. A secondary PCR was performed and clones for CC full length Cry j 1, pC190C91a and pC190C91d (identical clones) CC were identified. Cry j 1 or an antigenic fragment of it may be used CC for detecting, treating and preventing an allergic response to the CC Japanese cedar pollen allergen. It is capable of modifying both the CC B and T cell response to Cry j 1 and T cell response to a Cry j 1 CC antigen. The gene may be used to identify related allergens in other CC plant species.
 CC See also Q35305-20.
 CC Sequence 1337 BP; 408 A; 236 C; 292 G; 401 T;

alignment_scores: length: 27
 Quality: 71.00
 Ratio: 3.737
 Percent Similarity: 70.370 Percent Identity: 55.556

alignment_block:
 US-09-142-524-1 x Q35304 ..

Align seg 1/1 to: Q35304 from: 1 to: 1337

```

1 MetLysValThrValAlaIleAsnGlnPheGlyProAsnArgValPh 17
   |||||||.....:|||||:|||||:|||||:|||||:|||||
765 ATGAAGGTGACAGTGGCGCTTCAATCATTTGACCTAAGTGGACAAAG 814
17 eileLysArgValSerAsnValIleIleHis 27
   ::|||::|
815 AATGCCAGGCGACGATATGACTGTACAT 845

```

OM of: US-09-142-524-1 to: EST:* out_format : pfs

Date: Sep 29, 1999 1:50 PM

About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-1998 CompuGen Ltd.

Command line parameters:

-MODEL=frame+2n.model -DEV=xlp
-o=/cn2_1/USPTO_spool/US09142524/runat_29091999_125958_12773/app_query.fasta.1
-PB=EST -QFMT=fastap -SUFFIX=1st -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blonsum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR_SCORE=escore -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=stat -USER=US09142524 -NCPU=6 -ICPU=3 -WAIT
-THREADS=1

Search information block:

Query: US-09-142-524-1
Query length: 80
Database: EST:*
Database sequences: 2546578
Database length: 96626752
Search time (sec): 759.040000

Score list:

Score	Strd Orig	ZScore	EScore	Len	Documentation
9b_est17:AA559604	80.00	193.24	0.0300	519	AA659604 nt63e10.s1 NCI CGAP_P3
9b_est11:AA498109	77.00	185.30	0.0818	513	AA498109 v175n08.r1 Stratagene
9b_est11:AA198386	77.00	185.30	0.0831	520	AA198386 mw48e09.r1 Soares mouse
9b_est11:AA428242	74.00	180.84	0.1474	360	AA428242 csq1171.seq.f Human fe
9b_est129:AU036117	74.00	180.55	0.1528	371	AU036117 AU036117 Polyandrocarr
9b_est19:AA738025	75.00	180.03	0.1655	520	AA738025 nt15609.s1 NCI CGAP_GG
9b_est129:AU036104	73.00	179.30	0.1794	320	AU036104 AU036104 Polyandrocarr
9b_est12:AA300566	72.00	178.43	0.2006	265	AA300566 EST14320 Testis tumor
9b_est19:D76822	72.00	177.02	0.2404	308	D76822 MUR8BA01 mouse embryonal
9b_est11:AA603354	72.00	176.31	0.2633	583	AA603354 np06a11.s1 NCI CGAP_P3
9b_est11:AA732511	72.00	176.01	0.2736	343	AA732511 EST84502 Colon adenoca
9b_est11:AA381359	71.00	175.87	0.2788	268	AA381359 EST94431 Activated T-c
9b_est11:AA288847	71.00	175.69	0.2852	263	AA288847 m752c03.r1 Life Tech
9b_est11:AA378031	71.00	175.31	0.2996	490	AA378031 EST90680 Synovial sarc
9b_est11:AA815411	73.00	175.05	0.3097	287	AA815411 a161e03.s1 Soares test
9b_est13:AA349283	71.00	174.89	0.3099	504	AA349283 v175n08.r1 Life Tech
9b_est14:D55079	72.00	174.74	0.3175	293	AA349283 v175n08.r1 Life Tech
9b_est14:AA381326	71.00	174.73	0.3228	293	AA381326 EST94471 Activated T-c
9b_est14:AA386415	71.00	174.69	0.3241	298	AA386415 EST94471 Activated T-c
9b_est19:AA792506	73.00	174.48	0.3329	230	AA792506 v175n08.r1 Knowles SOL
9b_est12:AA300816	73.00	174.48	0.3330	535	AA300816 v175n08.r1 Knowles SOL
9b_est12:AA815996	73.00	174.46	0.3332	305	AA815996 v175n08.r1 Knowles SOL
9b_est10:AA815917	73.00	174.43	0.3338	536	AA815917 v175n08.r1 Knowles SOL
9b_est15:AA51865	73.00	174.32	0.3398	543	AA51865 v175n08.r1 Knowles SOL
9b_est13:AA338932	71.00	174.32	0.3398	543	AA338932 v175n08.r1 Knowles SOL
9b_est13:AA379634	71.00	174.20	0.3451	314	AA379634 EST92788 Skin tumor I
9b_est13:AA330864	71.00	173.97	0.3504	318	AA330864 EST92788 Skin tumor I
9b_est13:AA334566	71.00	173.94	0.3556	243	AA334566 v175n08.r1 Knowles SOL
9b_est13:AA33827	71.00	173.91	0.3570	323	AA33827 v175n08.r1 Knowles SOL
9b_est13:AA376323	70.00	173.78	0.3644	324	AA376323 v175n08.r1 Knowles SOL
9b_est12:AA321646	70.00	173.55	0.3650	329	AA321646 v175n08.r1 Knowles SOL
9b_est13:AA53470	70.00	173.55	0.3751	254	AA53470 v175n08.r1 Knowles SOL
9b_est13:AA381322	71.00	173.54	0.3757	337	AA381322 v175n08.r1 Knowles SOL
9b_est13:AA379368	71.00	173.49	0.3786	339	AA379368 v175n08.r1 Knowles SOL
9b_est13:AA360159	70.00	173.48	0.3786	256	AA360159 v175n08.r1 Knowles SOL
9b_est13:AA353178	70.00	173.44	0.3804	257	AA353178 v175n08.r1 Knowles SOL

9b_est13:AA377911 + 71.00 173.43 0.3811 341 ! AA377911 EST90625 Synovial s
9b_est13:AA379646 + 71.00 173.43 0.3811 341 ! AA379646 EST92511 Skin tumor
9b_est14:AA381374 + 71.00 173.35 0.3852 344 ! AA381374 EST94447 Activated

seq_name: gb_est17:AA659604

seq_documentation_block:

LOCUS AA659604 519 bp mRNA EST 05-NOV-1997
DEFINITION nt63e10.s1 NCI CGAP_P3 Homo sapiens cDNA clone IMAGE:1203210
SIMILAR to gb:X69150.40S RIBOSOMAL PROTEIN S18 (HUMAN); mRNA
SEQUENCE
AA659604
NID g2595758
KEYWORDS AA659604.1 GI:2595758
EST.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 519)
NCI CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
Tumor Gene Index
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
On May 9, 1995 this sequence version replaced gi:802232.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuang,
M.D., Michael Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLM at:
www.bio.limn.gov/bdnp/image/image.html

Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 377.
Location/Qualifiers
1. 519

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1203210"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site: 1; NotI; Site: 2; EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNAse-treated, total cellular RNA obtained from
5,000-10,000 microdissected cells
histologically-determined to be fully malignant prostate
cancer cells. Double-stranded cDNA was ligated to EcoRI
adaptor-specific primer, and the resulting PCR product
subcloned into pAMP10 by the UDG-cloning method (Life
Technologies). Average insert size is 600 bp. NOTE: Not
directionally cloned. This library was constructed by
David Krizman."

BASE COUNT
ORIGIN
145 a 114 c 143 g 117 t

alignment_scores:

Quality: 80.00 Length: 74
Ratio: 2.000 Gaps: 2
Percent Similarity: 54.054 Percent Identity: 32.432

alignment_block:
US-09-142-524-1 x AA659604

```

/organism="Mus musculus"
/strain="Inbred CD-1"
/db_xref="taxon:10090"
/clone="IMAGE:91811"
/clone_1fp="Stratagene mouse testis (#937308)"
/sex="males"
/tissue_type="testis"
/dev_stage="10-12 week old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: testis; Vector: pBluescript SK-, Site_1

```

Contact: Marra M/Mouse EST Project
WashU-HMM Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LIND ; contact the
IMAE Consortium (info@image.lind.gov) for further information.
MGI:404160

Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 477.
Location/Qualifiers
source

1..520
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:558312"
/clone_lib="Soares mouse 3NME12 5"
/sex="unknown"
/tissue_type="fetus"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
/note="Organ: whole fetus; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 3', on total mouse RNA (provided by Minoru Ko, Wayne State Univ.); double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 137 a 127 c 153 g 103 t
ORIGIN

alignment_scores:
Quality: 77.00 Length: 88
Ratio: 1.638 Gaps: 2
Percent Similarity: 53.409 Percent Identity: 30.682

alignment_block:
US-09-142-524-1 x AA198386 ..

Align seg 1/1 to: AA198386 from: 1 to: 520

```

3 ValThValAlaphaennglnpHeGlyProasnArgValPheIlely 19
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10 GTGACCGCGCGCCATGTCTAGTACCTCGAGAACTTCACAGACATTTT 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
19 sArgValSerAsnValIleIleHISglYArgArgIleAspIlePheAlas 36
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 GCGAGTACCTCACACACACATCGATGCGCGGCGGAAATAGCCTTCGCCCA 109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
36 eRlyAsnPhHisLeuGlnLyAsnThrIleGlyThrGlyArg... 51
|||:|||||:|||||:|||||:|||||:|||||:|||||:
110 TCACTGCCATTAG.....GCGTGGGCGGAGATAT 141
52 .....IleSerLeuLySLeuThSerG1 59
|||||:|||||:|||||:|||||:|||||:|||||:
142 GCTCATGTGTTGAGAAACAGACATCGACCTCACCAAGAGGCTGG 191
|||||:|||||:|||||:|||||:|||||:|||||:
59 yLysIleAlaSerArgArgValAspGlyIleIleAlaIleAlaTyrgInaSnP 76
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
192 AAAAATCAAGGAGATGAGTGGACGAGTATCATCATCATGACGAAC 241
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
76 rAlaSerTrpLys 80
|||||:|||||:|||||:|||||:|||||:|||||:
242 CACGACAGTACAAAG 255
|||||:|||||:|||||:|||||:|||||:|||||:
seq_name: gb_est11:AA248242

```

seq_documentation_block:
LOCUS AA248242 360 bp mRNA EST 11-MAR-1997
DEFINITION cs9117.seq.F Human fetal heart, lambda ZAP Express Homo sapiens
CDNA 5', mRNA sequence.
ACCESSION AA248242
NID 91878895
VERSION AA248242.1 GI:1878895
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 360)
Liew/C.C.
CDNAs from human fetal heart (1997)
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1402306.

Contact: Liew CC
Department of Laboratory Medicine and Pathobiology
University of Toronto
Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
Tel: 4169788758
Fax: 4169785650
Email: liewcc@utcc.utoronto.ca
PCR Primers
FORWARD: 5' GCCAGCTGAAATTACCCCTACTAAAGG 3'
BACKWARD: 5' CCGATGATTTGTAATGACATCAGCTATAGGGCG 3'
Seq primer: 5' GAAATTAACCTCATCAAGG 3'.
Location/Qualifiers
1..360
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="907F10"
/clone_lib="Human fetal heart, lambda ZAP Express"
/lab_host="E. coli XL1-Blue"
/note="Vector: lambda ZAP Express; Site_1: EcoRI; Site_2: XhoI; mRNA was purified from human fetal hearts (8-10 weeks). cDNA was synthesized using a XhoI-oligo dT adaptor-primer. EcoRI adaptors were ligated, followed by digestion with XhoI, for directional cloning into predigested lambda ZAP Express."

BASE COUNT 102 a 84 c 100 g 74 t
ORIGIN

alignment_scores:
Quality: 74.00 Length: 79
Ratio: 1.805 Gaps: 2
Percent Similarity: 51.899 Percent Identity: 31.646

alignment_block:
US-09-142-524-1 x AA248242 ..

Align seg 1/1 to: AA248242 from: 1 to: 360

```

12 ProAsnArgArgValPheIleLySerValSerValIleIleHISgl 28
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 CCGAAGAGTCCAGCATATTTTCCAGTACTACACACACATCGATG 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 yArgArgIleAspIlePheAlaSerLyAsnPhHisLeuGlnLyAsnT 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
84 GCGCGGAAATAGCCTTTCATCATCGCATTAG..... 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
45 hrIleGlyThrGlyArg... 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 .....GCTGGGCGCGAAGTATGCTCATGTGTTGAGAAACAGAC 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
52 IleSerLeuLySLeuThSerGlyIleAlaSerArgValAspG1 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 ATGACCTCACCAAGAGGCGGAGAGTACTGAGAGAGAGTGGACCG 215
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
68 yIleIleAlaIleAlaTyrgInaSnProAlaSerTrpLys 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 TGTGATCCATTATGCAAGATTCACGCGCATGACAG 252
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
seq_name: gb_est29:A0036117

```

seq_documentation_block:
LOCUS A0036117 371 bp mRNA EST 26-MAR-1999
DEFINITION A0036117 Polyandrocarpa misakiensis white spot budding stage
Polyandrocarpa misakiensis cDNA, mRNA sequence.
ACCESSION A0036117

NID 94527078
 VERSION AU036117.1 GI:4527078
 KEYWORDS EST.
 SOURCE Polyandrocarpa misakiensis.
 ORGANISM Polyandrocarpa misakiensis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 Stolidobranchia; Styelidae; Polyandrocarpa.

REFERENCE
 1 (bases 1 to 371)
 Kawamura, K., Hayata, D., Fujiwara, S. and Yubisui, T.
 Serine protease inhibitors expressed in the process of budding of
 tunicates as revealed by EST analysis
 J. Biochem. (1998) In press
 On Jun 5, 1998 this sequence version replaced gi:3188168.

JOURNAL
 COMMENT

CONTACT: Kawamura K
 Faculty of Science
 Kochi University
 2-5-1, Akebono-cho, Kochi 780, Japan
 Tel: +81-888-44-8313
 Fax: +81-888-44-8313
 Email: kazukiecc.kochi-u.ac.jp.

FEATURES
 source
 1..371
 /organism="Polyandrocarpa misakiensis"
 /strain="white spot"
 /db_xref="taxon:7723"
 /clone_lib="Polyandrocarpa misakiensis white spot budding
 stage"
 /dev_stage="budding stage"
 /dev_stage="budding stage"

BASE COUNT 103 a 83 c 100 g 84 t 1 others

ORIGIN

alignment_scores:
 Quality: 74.00 Length: 85
 Ratio: 1.644 Gaps: 3
 Percent Similarity: 52.941 Percent Identity: 25.882

alignment_block:
 US-09-142-524-1 x AU036117 ..

Align seg 1/1 to: AU036117 from: 1 to: 371

```

6 AlaPheAsnGlnPheGlyProAsnArgValPheIleLysArgValSe 22
:::|||||
83 AGTTTCACACATTC.....TTTCGTATCAC 108
:::|||||
22 rAsnValIleIleHisGlyArgArgIleAspIlePheLaserLysAsn 39
:::|||||
109 GAACACGATATGATGAGCGAGAGATGCTACGCC..... 148
:::|||||
39 heHisLeuGlnLysAsnThrIleGlyThrGlyArgArgIleSer..... 53
:::|||||
149 .....ATGACTGCAGTCAAGGAGTGGAGAGAGATTCAGCAACTTG 190
:::|||||
54 .....LeuLysLeuThrSerGlyLysIleAl 62
:::|||||
191 ATCTGCTTGAGCCGATGTTGATCTCAGAAACGTGCCGAGAACTGAC 240
:::|||||
62 aSerArgValAspGlyIleIleAlaIleAlaIleIleIleIleIleIle 79
:::|||||
241 AGAGGATGAAATCGACAGAGTGTCTACGATCTGCAATCTCTGCCCACT 290
:::|||||
79 rPLys 80
:::|||||
291 ACAAG 295
```

seq_name: gb_ests19:AA738025

seq_documentation_block:
 LOCUS AA738025 520 bp mRNA EST 22-JAN-1998
 DEFINITION nx15d09.s1 NCI CGAP GC3 Homo sapiens cDNA clone IMAGE:1256177 3'
 similar to gb:X69150 40S RIBOSOMAL PROTEIN S18 (HUMAN);, mRNA

sequence.
 ACCESSION AA738025
 NID 92768782
 VERSION AA738025.1 GI:2768782
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 1 (bases 1 to 520)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

JOURNAL
 COMMENT

CONTACT: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/FLN at:
www.dio.llnl.gov/bbrp/image/image.html

Insert Length: 579 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 402.

FEATURES
 source
 1..520
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="20q13.2-q13.3"
 /clone_image="IMAGE:1256177"
 /clone_lib="NCI-CGAP_GC3"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /note="Vector: pTR73D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from 3 pooled
 germ cell tumors, and was then primed with a Not I
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pTR73
 vector. Library is not normalized. Library was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 107 a 145 c 125 g 141 t 2 others

ORIGIN

alignment_scores:
 Quality: 75.00 Length: 79
 Ratio: 1.829 Gaps: 2
 Percent Similarity: 51.899 Percent Identity: 31.646

alignment_block:
 US-09-142-524-1 x AA738025/rev ..

Align seg 1/1 to reverse of: AA738025 from: 1 to: 520

```

12 ProAsnArgValPheIleLysArgValSerAsnValIleIleHisGI 28
:::|||||
485 CCGGAAGGTCACGATATCTCGAGTACCAACACCAACATGATGAG 436
:::|||||
28 yArgArgIleAspIlePheIleSerLysAsnPheHisLeuGlnLysAsn 45
:::|||||
435 TCGCGGAAATGACCTTTCCTGCTACCTGCATTAAG..... 399
:::|||||
45 hrlIleGlyThrGlyArgArg..... 51
:::|||||
398 .....GGTGTGGCCGAGATATGCTCATGTGTGTGAGGAAGCAGAC 354
```

```

98 ylllellAAlAAtyGtClAAsnDProAlaseTrrp 79
          ::::::::::: ||||| :::::
250 AGTGTGACGATCATGCAGATCTCGCAGTAC 283
seq_name: gb_est12:AA300566

seq_documentation_block:
LOCUS      AA300566          265 bp      mRNA
DEFINITION EST13420 Testis tumor Homo sapiens CDNA 5' end similar to similar
            to ribosomal protein S18, mRNA sequence.
ACCESSION  AA300566
NID        91952919
VERSION    AA300566.1 GI:1952919
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 265)
            Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirnesh,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wal,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodex,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Mamaros,S.M., Merrick,J.M.,
            Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utechtack,T.R., Weidman,J.F., Li,Y.,
            Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.U., Ruben,S.M.,
            Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
            Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of CDNA sequence
            Nature 377 (6547 Suppl), 3-174 (1995)
COMMENT    On Sep 12, 1996 this sequence version replaced gi:1395394.

JOURNAL    Bioinformatics
MEDLINE    The Institute for Genomic Research
COMMENT    9712 Medical Center Drive, Rockville, MD 20850 USA
            TEL: 3018699056
            FAX: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tcdb/ngl/ngl.html)
            Seq primer: M13 Reverse
            Location/Qualifiers
                1. 265
                    /organism="Homo sapiens"
                    /db_xref="ATCC (inhost):192428"
                    /db_xref="taxon:9606"
                    /clone_lib="Testis tumor"
                    /sex="male"
                    /dev_stage="adult"
                    /note="Organ: testis; Vector: pBluescript SK-; Site_1:
                    EcorI; Site_2: XhoI"
BASE COUNT      70 a      61 c      68 g      62 t      4 others
ORIGIN
Alignment_scores:
Quality: 72.00      Length: 79
Ratio: 1.756      Gaps: 2
Percent Similarity: 51.899      Percent Identity: 30.380

```

alignment_block:
US-09-142-524-1 x AA300566 ..

Align seg 1/1 to: AA300566 from: 1 to: 265

```

12  ProAsnArgArgValPheIleLysArgValSerAsnValIleIleHisG1 28
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30  CCGAAAGATCCAGCATATTTTCGAGTACACACACACACATCGATG 79
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
28  YARGArgIleAspIlePheAlaSerLysAsnPhenIleuGlnLysAsnT 45
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80  GCGGCGGAAATAGCCTTGCATCCTCCTCATTAAG..... 116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
45  hTleGlyThrGlyArgArg..... 51
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
117  ....GGTGGGGCGAGATATGCTCATGTGTGAGGAAACACAGC 161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
52  IleSerLeuLysLeuThrSerGlyLysIleAlaSerArgArgValAspG1 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
162  ATTGACTCTACCAAGAGCGGCGGAGAACCTCAGAGAGTGGTGAACG 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68  YIleIleAlaIleArgGlnAsnProAlaSerTrpLys 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
212  TGTGATCACCATTATNCAGATCCAGCCAGTACAG 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

seq_name: gb_est9:D76822

seq_documentation_block: 308 bp mRNA EST 07-OCT-1997
LOCUS D76822
DEFINITION MUS78A01 mouse embryonal carcinoma cell line F9 Mus musculus cDNA
clone 78A01, mRNA sequence.

ACCESSION D76822
NID D76822
VERSION D76822.1 GI:1596557
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 308)
AUTHORS Nishiguchi, S., Sakuma, R., Nomura, M., Zou, Z., Jeonatsilavong, J.,
Joh, T., Yasunaga, T. and Shimada, K.
TITLE A catalogue of genes in mouse embryonal carcinoma F9 cells
identified with expressed sequence tags

JOURNAL J. Biochem. 119, 749-767 (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1405026.

FEATURES
SOURCE Location/Qualifiers
1..308
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_id="78A01"
/clone_lib="mouse embryonal carcinoma cell line F9"
RASF COUNT 82 a 74 c 91 g 57 t 4 others

alignment_scores:
Quality: 72.00 Length: 79
Ratio: 1.756 Gaps: 2
Percent Similarity: 51.899 Percent Identity: 30.380

alignment_block:
US-09-142-524-1 x D76822 ..
Align seg 1/1 to: D76822 from: 1 to: 308

```

12  ProAsnArgArgValPheIleLysArgValSerAsnValIleIleHisG1 28
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
37  CCTGGAAGATCCAGCATATTTTCGAGTACACACACACATCGATG 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
28  YARGArgIleAspIlePheAlaSerLysAsnPhenIleuGlnLysAsnT 45
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87  GCGGCGGAAATAGCCTTGCATCCTCCTCATTAAG..... 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
45  hTleGlyThrGlyArgArg..... 51
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
124  ....GGGTTGGGGCGAGATATGCTCATGTGTGAGNAGACGAGC 168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
52  IleSerLeuLysLeuThrSerGlyLysIleAlaSerArgArgValAspG1 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
169  ATCGACTCTACCAAGAGCGGCGGAGAACCTCAGAGAGTGGTGAACG 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68  YIleIleAlaIleArgGlnAsnProAlaSerTrpLys 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
219  AGTTATCACCATTATNCAGATCCAGCCAGCAGTACAG 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

seq_name: gb_est17:AA603354

seq_documentation_block: 583 bp mRNA EST 08-OCT-1997
LOCUS AA603354
DEFINITION np06411.s1 NCI-CGAP P-3 Homo sapiens cDNA clone IMAGE:1115516
similar to gb:X69150 40S RIBOSOMAL PROTEIN S18 (HUMAN); mRNA
sequence.

ACCESSION AA603354
NID 92437215
VERSION AA603354.1 GI:2437215
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 583)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394196.

FEATURES
SOURCE Location/Qualifiers
1..583
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="8"
/clone="IMAGE:1115516"
/clone_lib="NCI-CGAP_P3"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: PAMP10; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected cells
histologically determined to be fully malignant prostate
cancer cells. Double-stranded cDNA was ligated to EcoRI

Insert Length: 635 Std Error: 0.00
Seq primer: ~40ml3 fwd. ET from Amersham
High quality sequence stop: 437.

adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into PAMP10 by the U93-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

BASE COUNT 118 a 161 c 130 g 174 t
ORIGIN

alignment_scores:

Quality: 74.00 Length: 79
Ratio: 1.805 Gaps: 2
Percent Similarity: 51.899 Percent Identity: 31.646

alignment block:

US-09-142-524-1 x AA603354/rev ..

Align seg 1/1 to reverse of: AA603354 from: 1 to: 583

```

12 ProSnaRgArGvAlPheIlleLySArGvAlSeRAsnVAlleIlleHsG1 28
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
556 CCTGAAGAAGTCCAGCATATTGCGAGTACTCAACACCAACATCGATCG 507
28 YArGArGIlleAsPIllePheAlaSeRlYsAsnPhEhSlleuGlnLYsAsnT 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
506 GCGGCGGAATAATAGCCTTGCCATCATCGCATTAAG..... 470
45 hrIlleGlyThrGlyArGArG..... 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
469 .....GCTGTGGCGGAGATATGCTATGTTGTGGAGAAACAGAC 425
52 IlleSerLeuLYsLeuThrSerGlyLYsIlleAlaSerArGArGvAlAsp1 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
424 ATTGACCTCCACCAAGAGCGGAGGAGACTCGCTGAGATAGAGTGGAACG 375
68 YIlleIlleAlaAlATyGlnAsnProAlaSerTrpLYs 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
374 TGTGATCCACATTATGCAAGATCCAGCCACGATCAAG 338

```

seq_name: gb_est13:AA372511

seq documentation block:

LOCUS AA372511 343 bp mRNA EST 21-APR-1997

DEFINITION EST84502 Colon adenocarcinoma IV Homo sapiens cDNA 5' end similar to similar to ribosomal protein S18, mRNA sequence.

ACCESSION AA372511

NID 92024904

VERSION AA372511.1 GI:2024904

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

AUTHORS Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fullmer,R.A., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wei,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitchugh,W.M., Fritchman,J.L., Geoghegan,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Heblum,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-T., Maturo,S.M., Merrick,J.M., Moreno-Palenzuela,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudke,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferris,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisener,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M., and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns

JOURNAL based upon 83 million nucleotides of cDNA sequence
MEDLINE Nature 377 (6547 Suppl), 3-174 (1995)
COMMENT 96026280
On May 8, 1995 this sequence version replaced gi:800909.

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavage@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1..343

/organism="Homo sapiens"

/db_xref="ATCC (inhost):176944"

/db_xref="taxon:9606"

/clone_lib="Colon adenocarcinoma IV"

/dev_stage="adult"

/note="Organ: colon; Vector: pBluescript SK-; Site:1: EcoRI; Site:2: XhoI"

BASE COUNT 99 a 75 c 90 g 74 t 5 others

ORIGIN

alignment_scores:

Quality: 72.00 Length: 79

Ratio: 1.714 Gaps: 2

Percent Similarity: 53.165 Percent Identity: 30.380

alignment block:

US-09-142-524-1 x AA372511 ..

Align seg 1/1 to: AA372511 from: 1 to: 343

```

12 ProSnaRgArGvAlPheIlleLySArGvAlSeRAsnVAlleIlleHsG1 28
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 CCTGAAGAAGTCCAGCATATTGCGAGTACTCAACACCAACATCGATCG 81
28 YArGArGIlleAsPIllePheAlaSeRlYsAsnPhEhSlleuGlnLYsAsnT 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
82 GCGGCGGAATAATAGCCTTGCCATCATCGCATTAAG..... 118
45 hrIlleGlyThrGlyArGArG..... 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 .....GCTGTGGCGGAGATATGCTATGTTGTGGAGAAACAGAC 163
52 IlleSerLeuLYsLeuThrSerGlyLYsIlleAlaSerArGArGvAlAsp1 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 ATTGACCTCCACCAAGAGCGGAGGAGACTCTGAGATGAGGTGGAACG 213
68 YIlleIlleAlaAlATyGlnAsnProAlaSerTrpLYs 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214 TGTGATCCACATTATGCAAGATCCAGCCACGATCAAG 250

```

seq_name: gb_est14:AA381359

seq documentation block:

LOCUS AA381359 263 bp mRNA EST 21-APR-1997

DEFINITION EST94431 Activated T-cells I Homo sapiens cDNA 5' end similar to similar to ribosomal protein S18, mRNA sequence.

ACCESSION AA381359

NID 92033679

VERSION AA381359.1 GI:2033679

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

AUTHORS Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS

1 (bases 1 to 263)
 Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fulcher, R.A.,
 Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,
 White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C.,
 Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
 Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S.,
 Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
 Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
 Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
 Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,
 Small, K.V., Spriggs, T.A., Uterback, T.R., Weidman, J.F., Li, Y.,
 Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
 Dimke, D., Feng, D.-F., Ferrite, A., Fischer, C., Hastings, G.A.,
 He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
 Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H.,
 Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,
 Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,
 Fraser, C.M., and Venter, J.C.
 Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 On Sep 12, 1996 this sequence version replaced gi:1405101.

TITLE
 JOURNAL
 MEDLINE
 COMMENT

Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tadb/hgl/hgl.html>)
 Seq primer: M13 Reverse.

FEATURES
source

1. 263
 /organism="Homo sapiens"
 /db_xref="ATCC (inhost):185717"
 /db_xref="taxon:9606"
 /clone_lib="Activated T-cells 1"
 /dev_stage="adult"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 71 a 62 c 70 g 57 t 3 others
 ORIGIN

alignment_scores:
 Quality: 71.00 Length: 79
 Ratio: 1.732 Gaps: 2
 Percent Similarity: 51.899 Percent Identity: 30.380

alignment_block:
 US-09-142-524-1 x AA381359 ..

Align seg 1/1 to: AA381359 from: 1 to: 263

```

12  ProAaAGArGAlpAlpHeIleYArGAlSeArAsnValleIleHieGl 28
13  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
43  CCTGAAGAAGTTCACGATATTTTCAGAGTACACACCAACATGATGATG 92
28  YArGArGAlleAspIlePheAlaSerLysAsnPhelHisLeuGlnLysAnt 45
29  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
93  GCGGCGGAATAAGCCTTCCATCATGTCACATTAAG..... 129
45  hrIleGlyThrGlyArGArG..... 51
130  ....GCTGTGGCGCGAATATGCTCATGTGTGTGAGGAAGACAC 174
52  IlleSerLeuLysLeuThrSerGlyLysIleAlaSerArGArGValAspGl 68
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

175 ATGACCTCACCAGAGCGCGGAGAACTCAGATGAGGTGAGACG 224

68 YIleAlaAlaAlaTyrrGlnAsnProAlaSerTrpLys 80

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

225 TGTGATCACCATTATGCGAATCCACGCGCATACAG 261

seq_name: gb_est13:AA379609
 seq_documentation_block: 268 bp mRNA EST 21-APR-1997
 LOCUS AA379609
 DEFINITION EST92467 SK1n tumor 1 Homo sapiens cDNA 5' end similar to similar
 to ribosomal protein S18, mRNA sequence.

ACCESSION AA379609
 NID 42031927
 VERSION AA379609.1 GI:2031927
 KEYWORDS EST.

SOURCE

ORGANISM

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

REFERENCE
AUTHORS

1 (bases 1 to 268)
 Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fulcher, R.A.,
 Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,
 White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C.,
 Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
 Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S.,
 Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
 Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
 Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
 Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,
 Small, K.V., Spriggs, T.A., Uterback, T.R., Weidman, J.F., Li, Y.,
 Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
 Dimke, D., Feng, D.-F., Ferrite, A., Fischer, C., Hastings, G.A.,
 He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
 Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H.,
 Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,
 Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,
 Fraser, C.M., and Venter, J.C.
 Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 On Sep 12, 1996 this sequence version replaced gi:1394110.

TITLE
JOURNAL
MEDLINE
COMMENT

Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tadb/hgl/hgl.html>)
 Seq primer: M13 Reverse.

FEATURES
source

1. 268
 /organism="Homo sapiens"
 /db_xref="ATCC (inhost):183986"
 /db_xref="taxon:9606"
 /clone_lib="Skin tumor 1"
 /dev_stage="adult"
 /note="Organ: skin; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI"

BASE COUNT 72 a 65 c 71 g 57 t 3 others
 ORIGIN

alignment_scores:
 Quality: 71.00 Length: 79
 Ratio: 1.732 Gaps: 2
 Percent Similarity: 51.899 Percent Identity: 30.380

alignment_block:
US-09-142-524-1 x AA379609

Align seg 1/1 to: AA379609 from: 1 to: 268

```

12 ProAsnArgValPheIleIleValSerAsnValIleIleHisG1 28
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
40 CCTGAAGAGTCCAGCATATTTNGAGTACTACCAACACCATCGATGG 89
28 ValArgIleAspIlePheIleAspIleAspIleHisLeuGlnLysAsn 45
90 GCGGCGGAAATAGCCTTGGCATCATCGCATTAAG..... 126
45 hTleGlyThrGlyArg..... 51
127 .....GGTGTGGCGGAGATATGCTCATGTGCTTGAGAAACGAC 171
52 IleSerLeuLysLeuThrSerGlyLysIleAlaIleAspArgValAspG1 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
172 ATTGACCTCACCAGAGGCGCGAGACTCAGTGAAGATGAGTGGAGAC 221
68 yIleIleAlaIleAlaIleArgIleAsnProAlaIleSerIlePlys 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
222 TGTGATCAGCATATGCGAGATCAGCGCATCAAG 258

```

seq_name: gb_est12:AA288347

seq_documentation_block:

LOCUS AA288347 490 bp mRNA EST 14-APR-1997
DEFINITION m52c03.r1 life Tech mouse embryo 15 5dpc 10667012 Mus musculus
CDNA clone IMAGE:601060 5' similar to gb:M76763 Mus musculus
ribosomal protein (MOUSE);, mRNA sequence.

ACCESSION AA288347
NID 91938068
AA288347.1 GI:1938068
EST.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 490)

Marra, M., Hallier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMNI Mouse EST Project
Unpublished (1996)
On Sep 12, 1996 this sequence version replaced gi:1394706.

TITLE
JOURNAL
COMMENT

FEATURES

source
1.490
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/map="xg27.3-q28"
/clone_image="601060"
/clone_lib="Life Tech mouse embryo 15 5dpc 10667012"
/tissue_type="embryo"
/dev_stage="15.5dpc embryos"
/lab_host="DH10B"

/note="Organ: whole embryo; Vector: PCMV-SPORT2; Site: 1;
Salt; Site: 2; Note: Cloned unidirectionally. Primer:
Oligo dT. 15.5dpc embryos. PCMV-SPORT2 vector."
BASE COUNT 122 a 129 c 150 g 85 t

alignment_scores:

Quality: 73.00 Length: 87
Ratio: 1.587 Gaps: 2
Percent Similarity: 52.874 Percent Identity: 29.885

alignment_block:

US-09-142-524-1 x AA288347

Align seg 1/1 to: AA288347 from: 1 to: 490

```

4 ThrValAlaPheAsnGlnPheGlyProAsnArgValPheIleLysAr 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
21 ACCGCGCCCATGTCTCTAGATCCCTGGAAGTCCAGCAGCATTTTCG 70
20 GValSerAsnValIleIleHisGlyArgIleAspIlePheIleSerL 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
71 AGTACTCAACACCAACATCATGAGGCGGCGGAAATAGCCTTCCGATCA 120
37 ysaAsnPheHisLeuGlnLysAsnThrIleGlyThrGlyArg..... 51
121 CTGCGCATTAAG.....GGTGTGGCGGAGATATGCT 152
52 .....IleSerLeuLysLeuThrSerGlyLysIleAlaIleAspArgValAspG1 60
|||||:|||||:|||||:|||||:|||||:|||||:
153 CATGTGTGTGAGAAAGACAGACATGCAGCTCAACCAAGAGGCTGGAGA 202
60 sIleAlaSerArgArgValAspGlyIleIleAlaIleAlaIleArgIleAsnProA 77
|||||:|||||:|||||:|||||:|||||:|||||:
203 ACTCAGCGAGATGAGTGGAGCGAGTGTATCCATCATGATGAGAAACCCAC 252
77 lAserIlePlys 80
|||||:|||||:|||||:|||||:|||||:|||||:
253 GACAGTACCAAG 263

```

seq_name: gb_est13:AA378031

seq_documentation_block:

LOCUS AA378031 287 bp mRNA EST 21-APR-1997
DEFINITION EST90680 synovial sarcoma Homo sapiens CDNA 5' end similar to
similar to ribosomal protein S18, mRNA sequence.

ACCESSION AA378031
NID 92030349
AA378031.1 GI:2030349
EST.
KEYWORDS EST.
SOURCE human.
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 287)

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-wai, C.,
Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S.,
Glock, A., Gnehm, C.L., Hanna, M.C., Heddlom, E., Hinkle, P.S., Jr.,
Kelley, J.M., Kelley, J.C., Liu, L.-T., Marineros, S.M., Merrick, J.M.,
Moreno-Parlanques, R.F., McDonald, L.F., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,
Small, K.V., Spriggs, T.A., Uteback, T.R., Weidman, J.F., Li, Y.,
Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dinke, D., Feng, D.-F., Ferrite, A., Fischer, C., Hastings, E.J.,
He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H.,
Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,
Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseitline, W.A., Fields, C.,
Fraser, C.M. and Venter, J.C.

TITLE Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
 MEDLINE 96026280
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1395403.

Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423

Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/cdb/hgi/hgi.html>)
 Seq primer: M13 Reverse

FEATURES

Location/Qualifiers
 1..287
 /organism="Homo sapiens"
 /db_xref="ATCC (inhost):182414"
 /db_xref="taxon:9606"
 /map="809B05: 18: 18p11.21-18p11.23: 17q21"
 /clone.lib="Synovial sarcoma"
 /sex="female"
 /tissue_type="synovial membrane"
 /dev_stage="adult, 20 yrs"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 78 a 68 c 77 g 60 t 4 others
 ORIGIN

alignment_scores:
 Quality: 71.00 Length: 79
 Ratio: 1.732 Gaps: 2
 Percent Similarity: 51.899 Percent Identity: 30.380

alignment_block:
 US-09-142-524-1 x AA378031 ..

Align seg 1/1 to: AA378031 from: 1 to: 287

```

12 Proasnaargargvalpneilelysargvalserasnvallelehisgl 28
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
38 CCTGAAAGTCCAGCATATTTCGAGTACTCAACACACATCGATGg 87
28 YARGARGTlleaspilphealaserlysasnphenisleuglnlysasnt 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 GCGGCGGAATAATAGCCTTGGCCATCACTGCCATTAAg..... 124
45 hrlleGlyThrGlyArg..... 51
|||||:|||||:|||||:|||||:|||||:|||||:
125 .....GGTGTGGCGGAGATATGCTCATGTGTGTGAGGAAGCAGAC 169
52 lleSerleuLysLeuThrSerGlyLysllealaserargvalaspgl 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 ATTGACCTCACCAGAGGCGGAGAGACTCACTGAGATGAGGTGGAACg 219
68 yllellelaalaIatyrGlnasnProAlaserTrpLys 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 TGTGATCACCATTATGCAGATCCAGCCAGTACAG 256
```